

ABOUT THIS USER GUIDE

This user guide is a practical guide to using Hermes, the 3D visualiser provided with the GOLD Suite and SuperStar packages. It includes comprehensive instructions for making the most of the visualisation and structure manipulation features available in the program so that docking, and other protein/ligand modelling results can be easily interpreted.

Use the < and > navigational buttons above to move between pages of the user guide and the **TOC** and **Index** buttons to access the full table of contents and index. Additional online Hermes resources can be accessed by clicking on the links on the right hand side of any page.

GOLD Suite users will also have two additional software modules accessible from Hermes:

- GOLD, for protein-ligand docking.
- GoldMine, for the analysis and post-processing of docking results.

GOLD and GoldMine help is available via the links given on the right hand side of any page. Access to GOLD and GoldMine is restricted in accordance with your current licence agreement.

Links to pre-computed libraries of intra and intermolecular geometry (Mogul and IsoStar respectively) are provided from Hermes to enable validation of ligand geometries and H-bond and non-bonded interactions. Mogul and IsoStar help is available via the links given on the right hand side of any page. Access to Mogul and IsoStar requires a licence for the Cambridge Structural Database System.

SuperStar licencees are able to access a SuperStar interface from within Hermes. Help is available via the links given on the right hand side of any page. Access to SuperStar is restricted in accordance with your current licence agreement. Access to SuperStar also requires a licence to the Cambridge Structural Database System.

The visualisation features in Hermes can be utilised without registration no matter which CCDC products are registered at your site. Hermes visualisation features can therefore be installed and used at the site licensed, e.g. to enable sharing of docking results between molecular modellers and bench chemists.

A tutorial is available for Hermes. GOLD and GoldMine tutorials are also available: these also illustrate some Hermes functionality. Tutorials can be accessed by clicking on the **Tutorials** link on the right hand side of any page.

The Hermes User Guide is divided into the following sections:

- 1 Overview of the Hermes Visualiser (see page 4)**
- 2 Introduction to Using the Hermes Visualiser (see page 5)**
- 3 Loading Structures, and Hits from Other Programs (see page 9)**
- 4 Navigating Proteins with the Protein Explorer Window (see page 11)**
- 5 Moving the Contents of the 3D Display Area (see page 19)**
- 6 Selecting and Deselecting Atoms and Molecules (see page 23)**
- 7 Customising the Display (see page 29)**
- 8 Labelling Atoms (see page 37)**
- 9 Calculating and Displaying Ribbons, Tubes and Other Graphics Objects (see page 41)**
- 10 Overlaying or Superimposing Proteins (see page 51)**
- 11 Measuring Geometrical Parameters (see page 55)**
- 12 Finding Hydrogen Bonds and Other Nonbonded Contacts (see page 57)**
- 13 Visualising CSD- and PDB-based Intramolecular Contact Distributions (see page 61)**
- 14 Validating Ligand Geometry (see page 65)**
- 15 Viewing Structures in Stereo (see page 73)**
- 16 Defining Descriptors (see page 75)**
- 17 Calculating and Using Descriptors (see page 91)**
- 18 Editing Structures (see page 93)**
- 19 Saving Files (see page 97)**
- 20 Appendix A: Special Details for SuperStar Users (see page 99)**
- 21 Appendix B: Special Details for GOLD Suite Users (see page 103)**
- 23 Appendix D: Calculating Descriptors: Computational Details (see page 117)**
- 24 Appendix E: Using Descriptors Tutorial (see page 121)**

CONDITIONS OF USE

The GOLD suite of programs (the "Program") comprising all or some of the following: Hermes (including as Relibase+ client and as SuperStar interface), GOLD, GoldMine, associated documentation and software, are copyright works of CCDC Software Limited and its licensors and all rights are protected. Use of the Program is permitted solely in accordance with a valid Software Licence Agreement or a valid Licence and Support Agreement with CCDC Software Limited or a valid Licence of Access to the CSD System with CCDC and the Program is proprietary. All persons accessing the Program should make themselves aware of the conditions contained in the Software Licence Agreement or Licence and Support Agreement or Licence of Access Agreement.

In particular:

- The Program is to be treated as confidential and may NOT be disclosed or re-distributed in any form, in whole or in part, to any third party.
- No representations, warranties, or liabilities are expressed or implied in the supply of the Program by CCDC Software Ltd., its servants or agents, except where such exclusion or limitation is prohibited, void or unenforceable under governing law.

GOLD © 2011 CCDC Software Ltd.
Hermes © 2011 CCDC Software Ltd.
GoldMine © 2011 CCDC Software Ltd.

Implementation of ChemScore, Heme, Kinase and Astex Statistical Potential scoring functions and the Diverse Solutions code within GOLD © 2001-2011 Astex Therapeutics Ltd.

All rights reserved

Licences may be obtained from:

CCDC Software Ltd.
12 Union Road
Cambridge CB2 1EZ
United Kingdom

Web: www.ccdc.cam.ac.uk
Telephone: +44-1223-336408
Email: admin@ccdc.cam.ac.uk

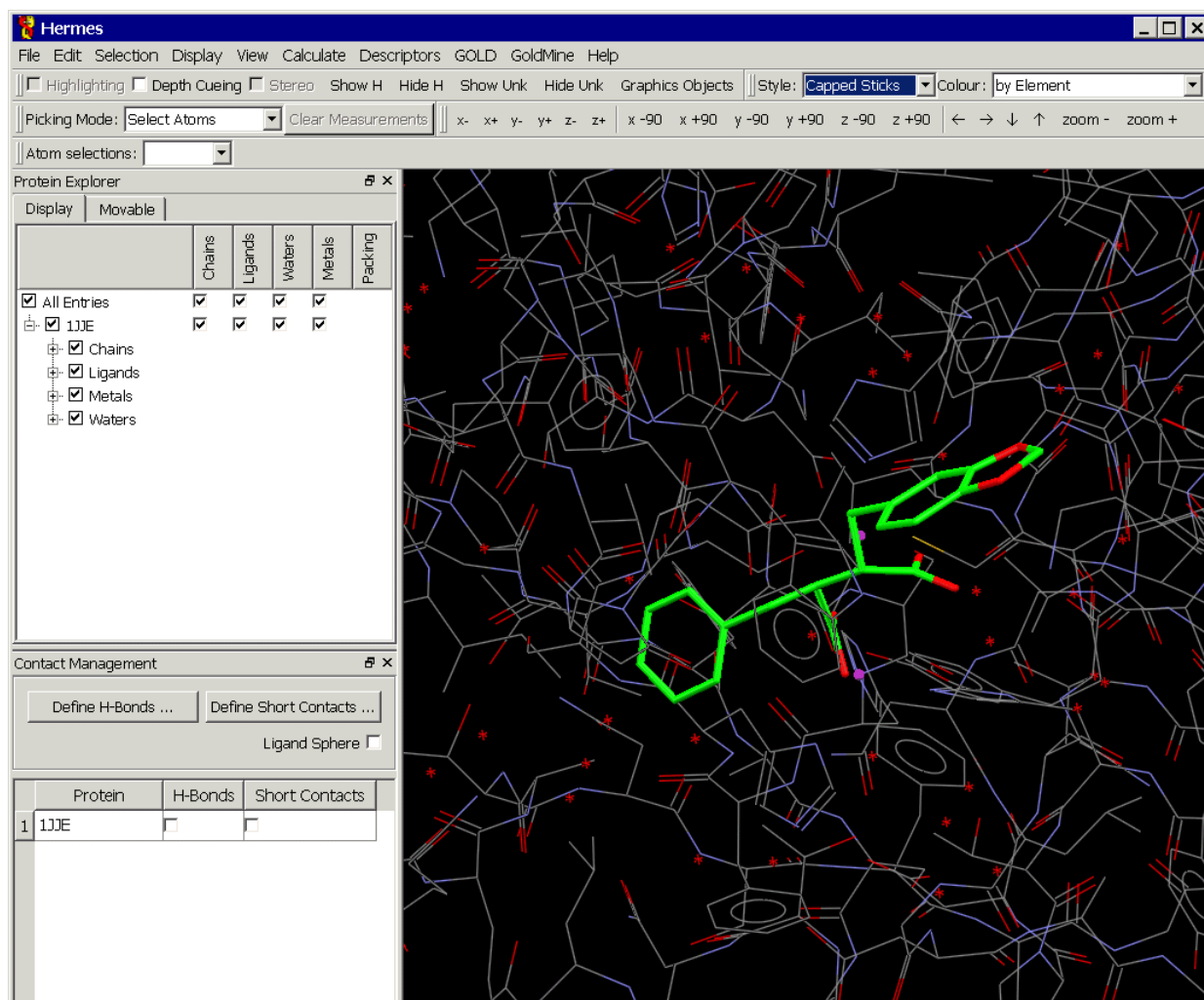
1 Overview of the Hermes Visualiser

The Hermes visualiser is a program for visualising protein structures in three dimensions. Its features include:

- The ability to read in protein and ligand structures from external files.
- The ability to rotate, translate and scale structures in the 3D display area.
- Very flexible functionality for navigating the protein structure and controlling what is displayed.
- A range of 3D visualisation options, including choice of display styles (wireframe, capped stick, ball and stick, and spacefill), colours, labelling schemes, and the ability to hide and then re-display atoms, residues, ligands, water molecules, etc.
- The ability to calculate and display auxiliary objects such as centroids and ribbons.
- The ability to overlay protein structures either by least squares overlay or by sequence alignment.
- The ability to measure and display distances, angles and torsion angles.
- The ability to find and display hydrogen bonds and nonbonded clashes, and to customise how they are defined geometrically.
- The ability to save displays.
- The ability to undo (reverse) previous action(s), re-apply these actions, and reset the display to default views.
- The ability to edit ligands, both manually and automatically.
- Links to Mogul to validate ligand geometries and IsoStar to inspect preferred locations and orientations of non-bonded interactions. Note that you must be a CSD System subscriber to access these features.
- The ability to load and visualise contoured surfaces.
- The ability to calculate descriptors.
- Z-clipping functionality.
- The ability to view structures in stereo.

2 Introduction to Using the Hermes Visualiser

2.1 Overview of the Interface



The visualiser interface consists of the following:

- *Top-level menu.*
- *Tool bar* containing common, basic options, e.g. for setting display styles and picking modes, and for flipping, rotating, translating and scaling the display.
- *Display area* for showing 3D structures.
- *Protein Explorer* window, for detailed control over what is displayed (may not be visible when Hermes is first opened).
- *Contact Management* window, for investigating nonbonded interactions (may not be visible when Hermes is first opened).

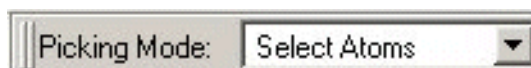
2.2 Top-Level Menu

- **File** allows you to read in proteins and ligands for display. It also provides facilities for saving what is displayed in the visualiser (see Saving Files, page 97). A recent files menu is also provided.
- **Edit** allows you to modify ligand structure, either manually or automatically (see Editing Structures, page 93). An option is also available for correcting ligand connectivity.
- **Selection** allows you to select sets of atoms in order for you to perform an action on them, such as colouring (see Selecting and Deselecting Atoms and Molecules, page 23).
- **Display** provides access to the following functionality:
 - Setting **Styles, Colours, Labels and Show/Hide** status (i.e. whether objects are displayed or hidden) (see Customising the Display, page 29).
 - Showing protein folding as a ribbon or tube display (see Creating Ribbon and Tube Displays, page 43).
 - Controlling general display preferences (see Setting General Display Properties, page 29).
 - Loading contoured surfaces (see Visualising Contoured Surfaces, page 45).
 - A dialogue for specifying IsoStar settings (see Setting Up Hermes For Use With IsoStar, page 61).
 - Loading of the *Graphics Object Explorer* window, which allows greater control over non-atomic graphics objects such as centroids, ribbon display, contoured surface display etc (see Controlling the Display of Centroids, Ribbons and Contoured Surfaces: the Graphics Objects Explorer, page 48).
- **View** allows you to open or close:
 - The *Protein Explorer window*, which provides options for navigating protein structures, e.g. finding particular residues easily (see Navigating Proteins with the Protein Explorer Window, page 11).
 - The *Contacts Window*, which provides functionality for finding hydrogen bonds and other short nonbonded interactions (see Using the Contacts Window, page 57).
 - The *Visualisation Options Toolbar*, which controls the display of H atoms, unknown atoms, whether or not depth cueing is on/off, whether or not the *Graphic Objects Explorer* is displayed etc.
 - *Docking Solution Information*. This is only relevant to GOLD Suite users and licence holders, please refer to the GOLD documentation (or (see Managing GoldMine Database Data: the GoldMine Pane, page 106)).
- **Calculate** provides access to the following functionality:
 - Creating and displaying centroids (see Creating Centroids, page 41).
 - Mogul geometry check (see Validating Ligand Geometry, page 65).
 - IsoStar library of intermolecular geometry (see Visualising CSD- and PDB-based

Intramolecular Contact Distributions, page 61).

- SuperStar for predicting interaction hotspots in protein active sites and around small molecules, see separate documentation for further details. Note that you must have a valid SuperStar licence to use this feature.
- The ability to overlay protein structures (see Overlaying Structures, page 51).
- The ability to superimpose protein structures e.g. to carry out an ensemble docking (see Superimposing Protein Structures, page 52).
- **Descriptors** provides access to the following functionality:
 - Definition of both simple and more complex descriptors (see Defining Descriptors, page 75).
 - Calculation of descriptors, once they have been defined ((see Calculating and Using Descriptors, page 91).
 - Export of the descriptor configuration file (see Saving Descriptor Definitions, page 89).
 - Activation of the *Descriptor Browser* tab within the *Protein Explorer* window. This can be used to select, edit and delete individual descriptors.
- **GOLD** launches the GOLD interface (assuming Hermes can find a valid GOLD Suite licence) enabling new protein-ligand dockings to be set-up, or re-loading of existing configuration files. Hermes features specific to GOLD users are provided (see Appendix B: Special Details for GOLD Suite Users, page 103). Comprehensive details on how to use the GOLD interface are provided in the GOLD documentation. From this menu option it is also possible to:
 - Load a gold fitting points file.
 - View and manage per atom scores (if this option has been switched on in GOLD).
- **GoldMine** provides a list of available options for GOLD users who have registered to use GoldMine. A valid GOLD Suite licence is required. GoldMine is a tool for analysing large quantities of docking information such as are produced from a structure-based virtual screen.
- **Help** provides access to on-line help, tutorials and registration information.

2.3 Picking Modes



The *Picking Mode* box in the tool bar controls what happens when you left-click on items in the display area:

- Use of **Select Atoms** mode permits the selection of atoms by clicking on them in the display area. Once selected, a set of atoms can be subjected to an operation, e.g. colouring, or used for a calculation of a centroid.
- Use of **Measure Distance**, **Measure Angle** or **Measure Torsion** modes permits the measurement of geometrical parameters by picking two, three or four atoms, respectively.

2.4 Right-Clicking

- Right-clicking in the 3D display area generates menus which provide access to many display options. The options available will depend on whether you right-click on:
 - A blank area in the display window, away from objects such as atoms, bonds and centroids.
 - Specific objects in the display, i.e. an atom, bond or ribbon.
- Right-clicking in other parts of the interface, e.g. the *Protein Explorer* window, will often display other context-sensitive pull-down menus.
- Menu items will be greyed out if they are inapplicable given the current state of the display.

2.5 Resetting the Display

- The display area can be returned to the default view direction, scale and perspective by right-clicking in the display background, selecting **View** from the resulting menu and then hitting **Reset View** in the next menu.
- To remove all nonbonded, distance, angle and torsion measurements, select **Clear Measurements** in the toolbar.

2.6 Several Ways of Doing Things

- There are often several ways of doing the same thing. For example, you may set the colour of centroids via the **Calculate** menu's **Centroids** button, or via the **Display** menu's **Colours** sub-menu, or by right-clicking on the centroid and using the resulting pull-down menu.

3 Loading Structures, and Hits from Other Programs

3.1 Loading Structures and Files from Other Programs

- Hermes is a hub for several CCDC products, specifically SuperStar, GOLD, GoldMine and Relibase+ and as such Hermes is diverse in terms of the type of structures and files it can read/generate.
- Hermes may also be used to display structures and sets of structures obtained using third party packages such as Sybyl and InsightII.
- Depending on which other software products you have a valid licence for, this may result in Hermes generating, displaying or being able to load:
 - Scatterplots and contour maps.
 - GOLD configuration files (files that store the setup for a protein-ligand docking).
 - GOLD docking results.
 - GoldMine database files (to enable post-processing of GOLD docking results).
 - Individual proteins or ligand binding sites.
 - Superimposed binding sites.
 - Superimposed hits from a 3D substructure search.

3.2 Loading a Structure from a File

- The Hermes visualiser has its own file format (filename extension **.reliarch*).
- The **.reliarch* file is binary, so cannot be read using a text editor.
- In addition, the visualiser can read structures from the following types of files:
 - **.cif*.
 - **.mol2*.
 - **.pdb*.
 - **.acnt* and **.omap* (see Visualising Contoured Surfaces, page 45).
- A file in any format known to the Hermes visualiser can be loaded in several ways:
 - Hit **File** in the top-level menu, and then select **Open**. Choose the file type (i.e. format) you want and then select the file you wish to open.
 - If you are using a computer that supports *Drag and Drop* (such as Windows 98 or later), you can drag a file icon and drop it onto the the Hermes program (or a shortcut to it). Hermes will launch and open the dropped file.
 - If you are using a computer that supports *Drag and Drop* (such as Windows 98 or later), and Hermes is already running on your computer, you can drag a file into the the Hermes visualiser window, and Hermes will open the file.
- The **Recent Files** menu details any recently loaded files: select the relevant file from the list to

load it.

3.3 Deleting Structures

- To delete a structure, open the *Protein Explorer* (see Opening the Protein Explorer Window, page 11), right-click on the identifier of the protein you wish to delete, and select **Delete** from the resulting pull-down menu. Components of a protein structure (e.g. specific chains, ligands, metals or water molecules) may also be deleted in the same way.
- To delete all structures, pick **File** from the top-level menu and **Close All Files** from the resulting pull-down menu.

4 Navigating Proteins with the *Protein Explorer* Window

4.1 The Main Purpose of the *Protein Explorer* Window

Protein structures are inherently hierarchical: proteins consist of chains, ligands, etc., chains consist of residues, residues consist of atoms. It is often necessary to locate particular parts of a protein (e.g. particular residues, ligands or atoms) in a complex 3D display. The *Protein Explorer* window aids this by presenting proteins in a tree layout, allowing specific groups of atoms to be located and selected easily.

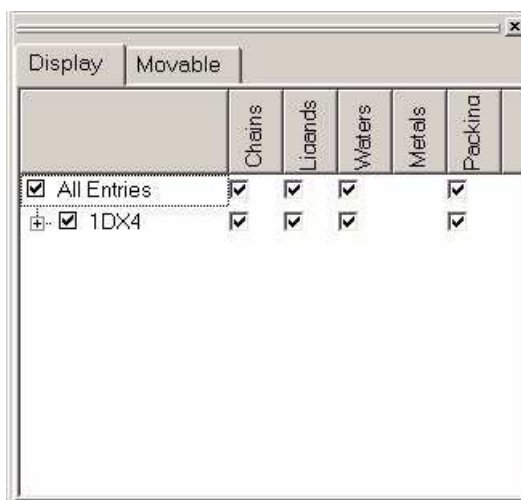
4.2 Opening the *Protein Explorer* Window

- The *Protein Explorer* is opened by selecting **View** from the top-level menu followed by **Protein Explorer** in the resulting pull-down menu (if there is a tick opposite **Protein Explorer** in this menu, the window is already open).
- The *Protein Explorer* is a *dockable window*, which means it can be placed in various positions in the main Hermes visualiser window, or kept as an entirely separate window. To do this, move the mouse cursor to the top of the *Protein Explorer* window so that the four-arrow icon appears. Then drag the window with the mouse, keeping the left-button depressed, and put it down where you want by releasing the mouse button.

4.3 Using the *Protein Explorer* Window

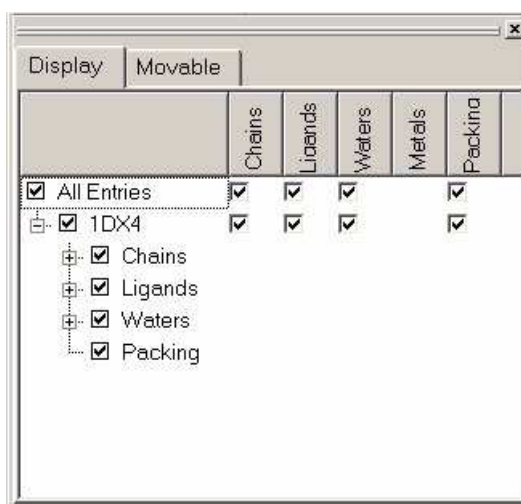
4.3.1 Overview of the *Protein Explorer Display Layout*

- When first opened, the *Protein Explorer* lists the identifiers of all loaded proteins, e.g.

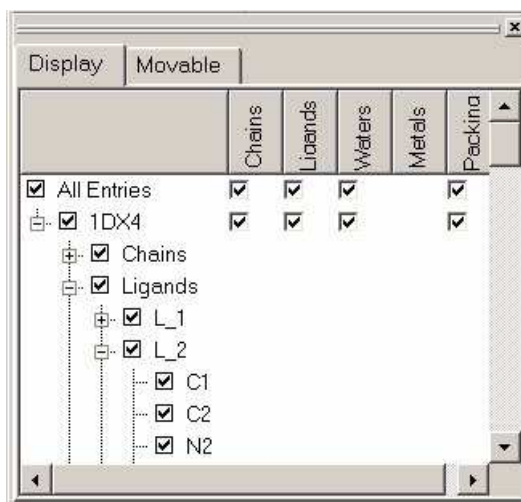


- In this screenshot, the *Display* tab is selected; the functionality of the other tabbed pane, *Movable*, is described elsewhere (see Moving Structures Independently of One Another, page 20).

- Two other tabbed panes can also be brought up in the *Protein Explorer Display*. These are the *Docking Solutions window* and the *Descriptor browser window* (see Displaying, Editing and Deleting Descriptor Definitions, page 89).
- Clicking on the + icon next to a protein identifier lists the major elements comprising the protein structure, viz. polypeptide chains and (if any are present) bound ligands, water molecules, metal ions and special covalent bonds (disulphide bridges and protein-ligand covalent linkages), e.g.

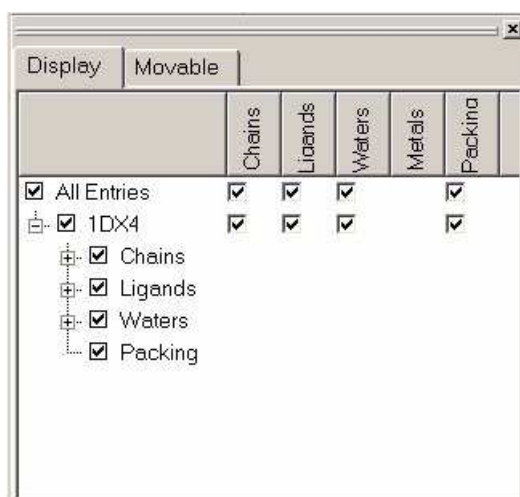


- Successively clicking on further + icons enables you to drill down to smaller and smaller entities, e.g. individual protein-chain residues and individual ligand atoms:

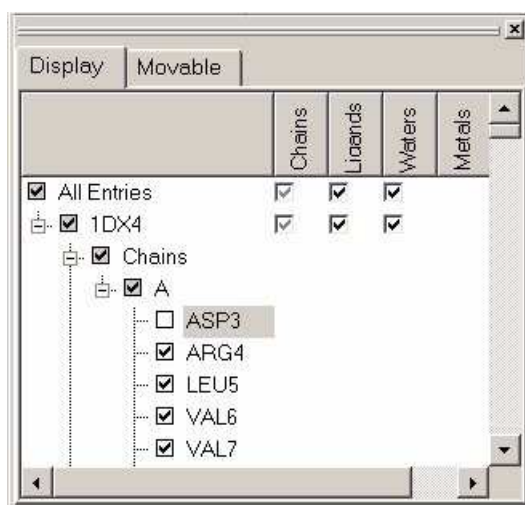


4.3.2 Using the *Protein Explorer* to Control What is Displayed

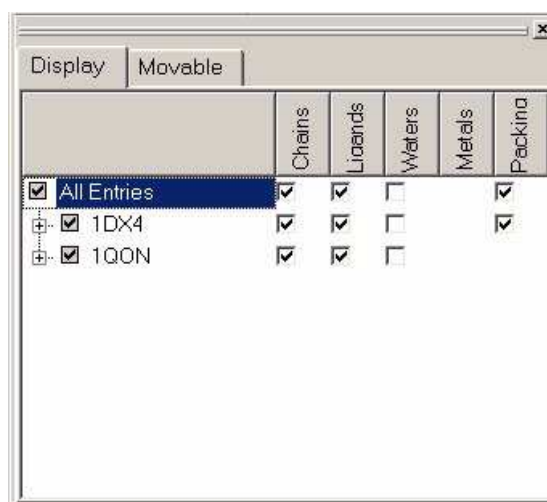
- There is a check box immediately to the left of each item listed in the *Protein Explorer Display* pane:



- Switching a check box on (as in the above screenshot) causes the corresponding atoms and bonds to be displayed in the 3D display area; switching it off causes them to be undisplayed.
- There is a complication if some but not all items at a particular level of the tree are displayed. In the example below, *ASP3* of chain *A* is switched off (not displayed in the 3D visualiser) but all other residues are switched on:



- The check box for chain *A* is therefore shown with a tick on a grey background, indicating that the chain is partially displayed. If this check box is successively clicked, it will cycle through the sequence: everything in the chain displayed; nothing displayed; partial display.
- The check boxes in the very top row of the *Protein Explorer Display* pane apply to all entries, e.g. to turn off the display of water molecules for all currently-loaded proteins, set the check boxes in the top row to:

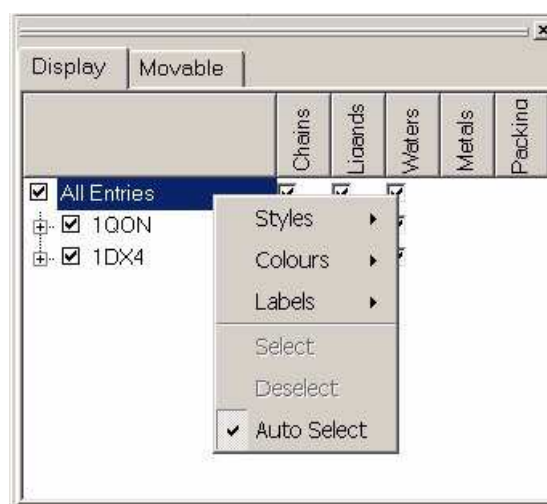


4.3.3 Using the *Protein Explorer* to Make Atom Selections

There are two possible ways of working, depending on whether *auto select* is switched on.

Auto Select On

- To switch auto select on, right-click in the *Protein Explorer Display* pane and pick **Auto Select** from the resulting pull-down menu (assuming this item is not already ticked, in which case it is switched on already):



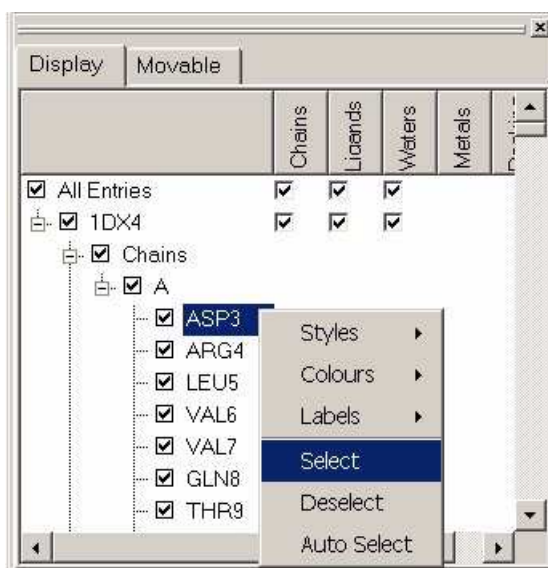
- Clicking on an item in the *Protein Explorer Display* pane will now select the corresponding atoms in the 3D visualiser, e.g. clicking on a chain identifier will select all the atoms of that chain, clicking on a protein identifier will select all atoms in that protein structure.
- Clicking on a different item will then select the atoms corresponding to that item. The atoms corresponding to the first item will become deselected unless:
 - the *Ctrl* key is pressed while the mouse is clicked, in which case the atoms corresponding to

the first item will remain selected;

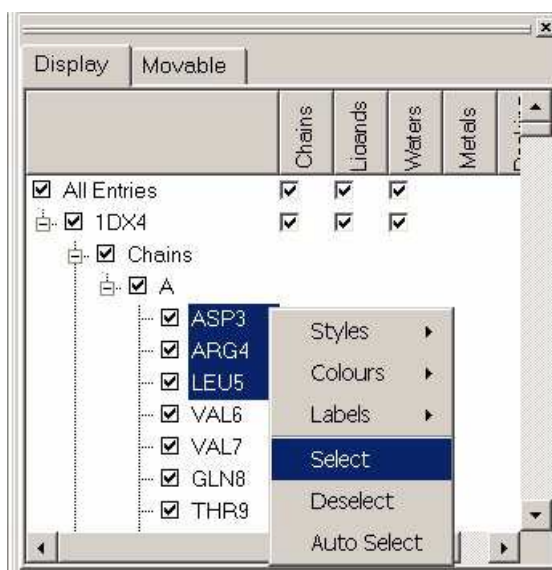
- the *Shift* key is depressed; this will select the atoms corresponding to the first item, the new item, and everything in between.
- Auto select is useful for finding things rapidly in the 3D display area. For example, if you click on *TRP84* in the *Protein Explorer*, that residue is immediately selected in the visualiser, and is therefore easy to locate. However, switching auto select on slows the program, so the option is turned off by default.

Auto Select Off

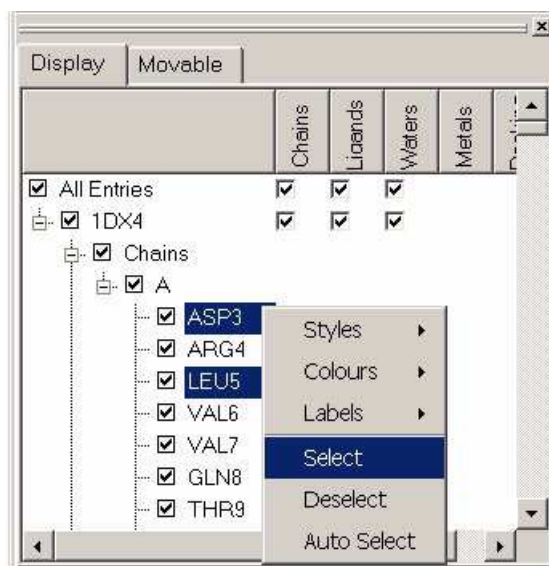
- If auto select is off, atoms in the 3D display area can be selected by right-clicking on an item in the *Protein Explorer* and picking **Select** in the resulting pull-down menu, e.g. to select the atoms of *ASP3*:



- To select atoms corresponding to a sequence of items in the *Protein Explorer Display* area, left-click on the first of the items, then left-click on the last of the items (or drag the cursor down to this item) with the *Shift* key depressed, then right-click anywhere in the highlighted (dark blue) area and pick **Select** from the resulting pull-down menu. For example, to select the atoms of *ASP3*, *ARG4* and *LEU5*:



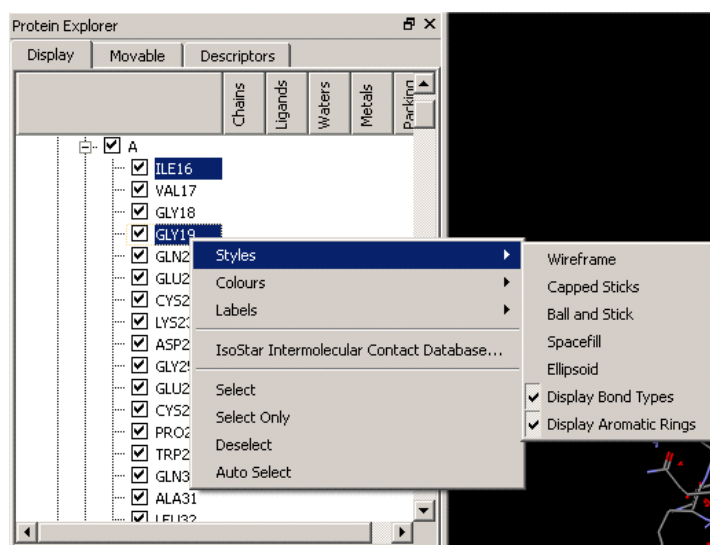
- Use of the *Ctrl* key allows non-sequential items to be picked by left clicking; the corresponding atoms can then be selected with the right-button pull-down menu. For example, to select the atoms of *ASP3* and *LEU5* but not *ARG4*:



4.3.4 Using the *Protein Explorer* to Change Display Properties

- Right-click on an item in the *Protein Explorer Display* pane and pick the required option (**Style**, **Colours**, etc.) from the resulting pull-down menu to change the display properties of the corresponding atoms in the 3D display area.
- To change properties for several parts of the protein simultaneously, select the required items in the *Protein Explorer*, using the *Ctrl* or *Shift* keys as necessary (see Using the Protein Explorer to Make Atom Selections, page 14).

- Then right-click anywhere in the highlighted (dark blue) part(s) of the *Protein Explorer Display* pane and pick the required option (**Style**, **Colours**, etc.) from the resulting pull-down menu, e.g.

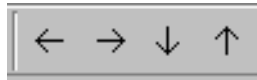


5 Moving the Contents of the 3D Display Area

5.1 Translating the Contents of the 3D Display Area

The contents of the display area can be translated in several ways:

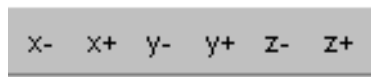
- By holding the middle mouse button down while moving the cursor in the display area (three-button mouse only).
- By moving the cursor in the display area while keeping both the left-hand mouse button and the keyboard *Ctrl* key pressed down.
- By using the arrow buttons in the tool bar:



5.2 Rotating the Contents of the 3D Display Area

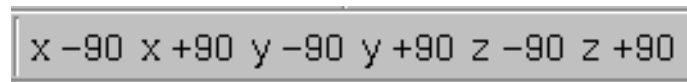
The contents of the display area can be rotated in several ways:

- By moving the cursor in the display area while keeping the left-hand mouse button pressed down (*x* and *y* rotation), or while keeping both the left-hand mouse button and the *Shift* key pressed down (*z* rotation).
- By using the buttons labelled *x-*, *y-*, etc. in the tool bar:



5.3 Flipping the Display by 90 Degrees

The contents of the display area can be rotated by 90 degrees about any of three orthogonal directions by using the buttons labelled *x-90*, *x+90*, etc. in the tool bar:



5.4 Scaling the Display

The contents of the display area can be scaled (i.e. zoomed in or out) in several ways:

- By moving the cursor up and down in the display area while keeping the right-hand mouse button pressed down.
- By using the buttons labelled **zoom-** and **zoom+** in the tool bar:

zoom - zoom +

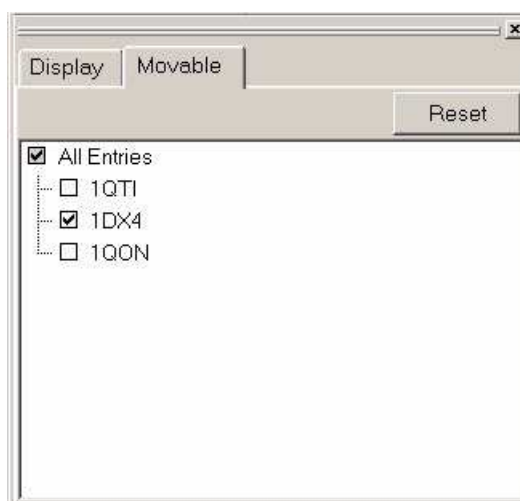
5.5 Changing the Display Perspective

The display perspective can be changed by moving the cursor up and down in the display area while keeping the right-hand mouse button and the *Shift* key pressed down. Like depth cueing (see Depth Cueing, page 30), perspective is a way of achieving a 3D effect, and it is unlikely to be helpful to use both at the same time (and, of the two, depth cueing is normally preferable).

5.6 Moving Structures Independently of One Another

If two or more structures are loaded, they can be moved independently of one another, as follows:

- Open the *Protein Explorer* (see Opening the Protein Explorer Window, page 11), if it is not already open.
- Select the tab labelled **Movable** in the Protein Explorer dialogue.
- All loaded structures will be listed in the *Protein Explorer*; use the check box next to each to determine whether it will move in response to mouse movement (check box ticked) or remain stationary, e.g. with the settings below:



1DX4 will move in response to the mouse but *1QTI* and *1QON* will not.

5.7 Setting the Rotation Centre

- By default, the centre of rotation will be at the centroid of all the loaded atoms.
- To set the rotation centre at a particular atom, right-click on the atom and select **Rotation Centre** followed by **Set Rotation Centre** in the resulting pull-down menus.
- To set the rotation centre back to its default, right-click anywhere in the display area and select

Rotation Centre followed by **Reset Rotation Centre**.

6 Selecting and Deselecting Atoms and Molecules

6.1 The Purpose of Selecting Atoms and Molecules

- Selection of atoms or molecules (e.g. ligands, protein chains) is useful for changing properties such as display style, calculating objects such as centroids, and for defining descriptors etc. Once a set of atoms has been *selected*, then one or more operations can be performed on those atoms.

6.2 Ways of Selecting Atoms and Molecules

Atoms may be selected or deselected in several ways:

- Hit **Selection** in the top-level menu to get access to **Select All**, **Deselect All** and **Invert Selection** buttons.
- Hit **Selection** in the top-level menu and then **Select** or **Deselect** in the pull-down menu. The resulting menu allows you to select or deselect all atoms of a particular type, e.g. **Non-Hydrogen**, **Backbone Atoms**.
- Hit **Selection** in the top-level menu and then **Select Molecules...** or **Select Atoms...** ; individual molecules or atoms, respectively, can then be selected by left-clicking on them.
- Right-click anywhere in the display area (atom, bond or background) and choose **Selection** from the resulting menu. This gives the same pull-down menus as described above.
- Set the *Picking Mode* (in the tool bar near the top of the main Hermes window) to **Select Atoms** and click on individual atoms with the left mouse button to select them. Once selected, an atom can be deselected by clicking on it again.
- To select all atoms that are directly bonded to atoms that have already been selected, right-click anywhere in the display area and choose **Selection** followed by **Expand Selection** from the resulting menu.
- Use the *Protein Explorer* to select chains, ligands, residues, etc., when you know the relevant identifier (e.g. the residue number) but you do not know where in the 3D display they are (see Using the Protein Explorer to Make Atom Selections, page 14).
- Complex selections of protein atoms can be made on the basis of property (e.g. Hydrophobic non-H atoms), atom type, proximity to the ligand etc. (see Making Complex Selections: Defining Subsets of Protein Atoms, page 24). Complex selections are required for the definition of descriptors (see Defining Descriptors, page 75).
- When the *Picking Mode* is set to **Select Atoms**, all atoms become deselected if you left-click anywhere in the display-area background.
- Press the *shift* key and then click on any atom to select or deselect the entire molecule containing that atom (when in **Select Atoms** mode).

6.3 Making Complex Selections: Defining Subsets of Protein Atoms

6.3.1 The Purpose of Protein Subsets

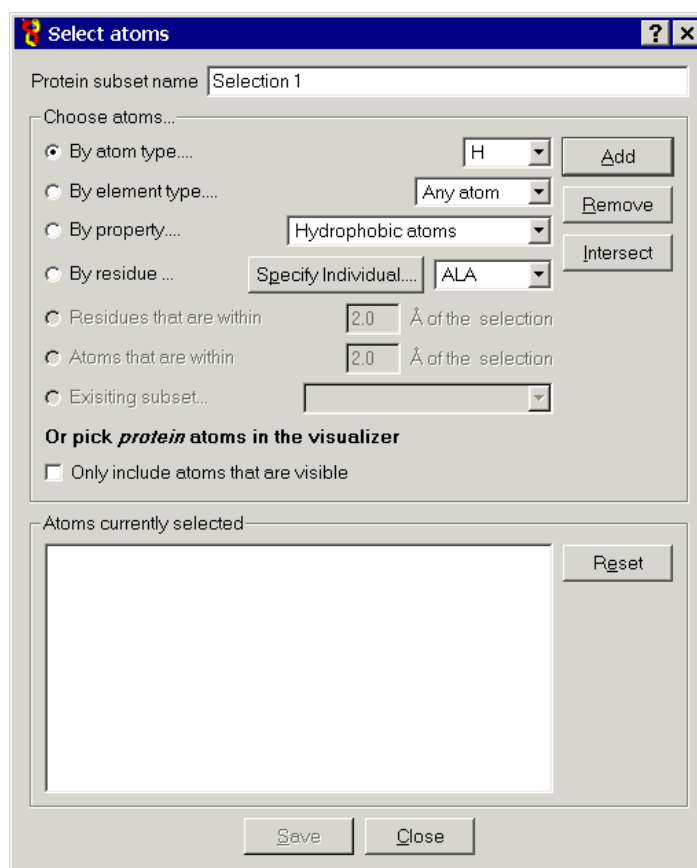
- Complex selections (i.e. *subsets*) of protein atoms are required for the definition of descriptors (see Defining Descriptors, page 75). For example:
 - if you want a descriptor that tells you whether NE2 of Gln42 forms a hydrogen-bond, then you need to create a protein subset containing just this one atom.
 - if you want a descriptor that tells about close contacts to protein aromatic carbons in the binding site, you need to create a protein subset containing all the aromatic carbons within, say, 5Å of one of the ligands (choose one of the larger ligands so that you get all relevant protein atoms).
- Creating complex selections is also a good way to simplify the visualiser display. For example, if you create a subset of protein residues within 5Å of the ligand, you can then hide all protein atoms that do not belong to this subset.

6.3.2 Creating Protein Subsets

- A complex atom selection is a subset of the atoms in the currently-loaded protein.
- A protein subset can be displayed and/or used in the definition of descriptors (see The Purpose of Protein Subsets, page 24).

Note: If you intend to use a subset in a hydrogen-bond descriptor, then you must ensure that the hydrogen atoms are included on all H-bond donor atoms. If a defined subset contains H-bond donor atoms but does not include all the hydrogen atoms to which they are bonded, then you will be given the option to add the required hydrogen atoms automatically.

- To create a protein subset, select **Selection** from the top-level menu and **Define Complex Selection...** from the resulting pull-down menu. This will open the *Select atoms in...* dialogue (this can also be opened from several of the descriptor-definition dialogue boxes):



- The white box at the bottom of the dialogue lists the protein atoms currently in the subset. This will be empty to begin with unless you have some atoms selected in the 3D visualiser when you open the protein-subset dialogue.
- To add atoms to the subset, switch on one of the radio buttons in the top part of the dialogue box (e.g. *By property...*), set the selection box alongside the button to the desired selection (e.g. *Hydrophobic non-H atoms*) and hit the **Add** button. If the *Only include atoms that are visible* check box is ticked, atoms that have been hidden in the visualiser will not be added to the subset when you do this.
- To remove atoms from the subset, follow the above procedure but hit **Remove** instead of **Add**. Hitting **Reset** will remove all atoms.
- To select some atoms and then intersect them (perform a logical AND) with the atoms currently in the subset, follow the above procedure but hit **Intersect** instead of **Add**.
- The *By atom type...* option allows you to select atoms according to their Sybyl atom type (see Atom Types, page 117).
- Atoms in particular residues can be added, removed or intersected by hitting the **Specify Individual** button and clicking on the check box adjacent to each required residue; then using **Add**, **Remove** or **Intersect** as desired.
- The *Residues that are within...* and *Atoms that are within...* options allow selection of residues or

atoms that lie within a specified distance of the current selection.

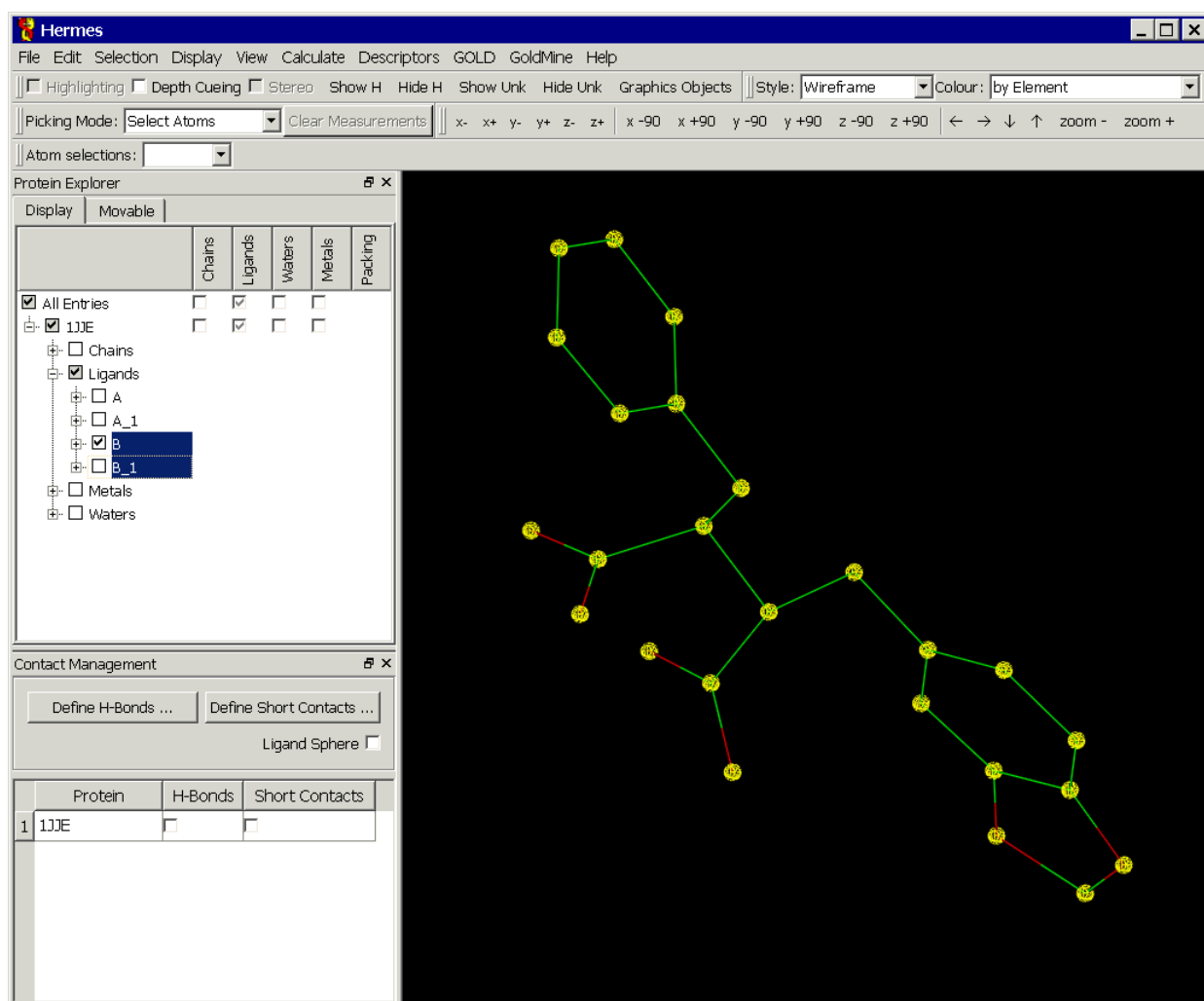
- Atoms can also be added to the subset by directly selecting them in the visualiser and can be removed by deselecting them.
- A default name will be supplied for the subset; this can be changed using the entry box at the top of the dialogue.
- Once you have included all the atoms you want in the subset, hit **Save** to define the subset. The subset is not created until you do this.
- Alternatively, those atoms currently selected in the display can be quickly saved as a protein subset by selecting **Selection** from the top-level menu and **Save Current Selection...** from the resulting pull-down menu.

6.3.3 Visualising Saved Protein Subsets

- If you have set up protein subsets (see Making Complex Selections: Defining Subsets of Protein Atoms, page 24) you can highlight the atoms belonging to any one of them by picking the required subset from the *Atom Selections* pull-down menu, which is situated above the visualiser display area.
- Information specific to saved protein subsets and other CCDC packages is provided elsewhere (see Visualising Saved Protein Subsets Within Hermes, page 109).

6.4 Identifying Selected Atoms

- Atoms which are selected are indicated by a network of yellow lines that wrap around the atom, e.g.



7 Customising the Display

7.1 Setting General Display Properties

7.1.1 Setting Lighting Effects

- The display area can be thought of as being illuminated by ambient lighting (i.e. low lighting from all directions) and up to four directional lights. Ambient lighting and two of the directional lights are switched on by default. (The lighting has no effect if you are displaying everything in wireframe style.)
- To control the lighting, hit **Display** in the top-level menu followed by **Display Options...** in the next menu. Ensure that you are on the *Lighting* pane of the resulting *Display Options* dialogue.
- Switch each light on or off by use of the *Enabled* check boxes.
- Each directional light lies along a vector direction which can be altered by typing in new values in the *Position* *x*, *y* and *z* entry boxes. (1,1,1) lies above the top-right of the screen, (-1,-1,-1) lies below the bottom-left.
- Any colour can be altered by clicking on the colour panel and selecting from the resulting palette (see Using the Select Color Dialogue Box, page 34). In the case of ambient lighting, the colour may be changed by clicking on the **Diffuse ambient colour:** panel and selecting from the colour palette.
- *Diffuse colour* refers to the general background lighting. *Specular colour* controls the colour of the reflections off atoms when displayed in space-filling mode. You can set the specular colour to black to remove these reflections completely.
- Hit the **Defaults** button to revert to default settings.

7.1.2 Setting Background Colours

- You can switch between the default black background and an alternative colour by hitting **Display** in the top-level menu followed by **Display Options...** in the next menu. Move to the *Background* pane of the resulting *Display Options* dialogue.
- To set a customised, uniform background colour, select the *Single Colour* option, click on the adjacent panel (which will be black by default) and choose a colour from the resulting palette (see Using the Select Color Dialogue Box, page 34).
- To create a background which changes gradually in colour, select the *Gradient* option, click on each of the four panels in turn and select colours from the resulting palette to achieve the effect of your choice.
- Select the *None (Black)* option or hit the **Defaults** button to revert to the default black background.

7.1.3 Depth Cueing

- Depth cueing aims to improve the 3D effect by displaying objects at the front more brightly than those at the back. The extent to which this improves 3D perception varies according to what is on display and it may not always be effective.
- To customise the depth-cue settings, hit **Display** in the top-level menu followed by **Display Options...** in the next menu. Move to the *Depth Cueing* pane of the resulting *Display Options* dialogue.
- Type in values in the *Front brightness* and *Back brightness* boxes to control how bright an object at the front is compared to one at the back. The numbers must fall in the range 0.0 to 1.0.
- Select the *Equation* to be used for calculating the depth cueing. **Linear** or **Exponential2** give better results when dealing with large structures.
- The *Enabled* check box can be used to switch depth cueing on and off.
- Hit the **Defaults** button to revert to default depth-cue settings.

7.1.4 Z-Clipping

- Z-clipping restricts the contents of the rendered image based on the Z coordinate of individual objects. There is both a *near clipping plane* and a *far clipping plane*. Objects closer to the viewer than the front plane are not rendered, and objects further from the viewer than the rear plane are not rendered.
- To customise the positions of the *near* and *far clipping planes*, hit **Display** in the top-level menu followed by **Display Options...** in the next menu. Move to the *Z-Clipping* pane of the resulting *Display Options* dialogue.
- The *Enabled* check box can be used to switch Z-Clipping on and off.
- The *Far Plane* and *Near Plane* sliders allow interactive control over the position of the clipping planes. At any given slider setting, the top and bottom white boxes show the limits of the current display. Slider settings can take the values from -100 to +100.
- You can type distance limits directly into the white boxes, but you must terminate each number with a carriage return (i.e. *Enter*) for it to take effect.
- Moving the middle grey rectangle with the mouse alters both sliders simultaneously. By setting a narrow range, this therefore allows you to view slices through the rendered image.
- Hit the **Defaults** button to revert to default Z-Clipping settings.

7.2 Setting Default Style and Colour Preferences for Different Parts of the Protein

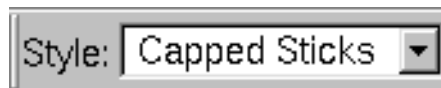
- To control the colours and styles that are applied by default to different parts of a protein when it is first loaded (peptide chains, ligands, water molecules, metal ions), hit **Display** in the top-level menu followed by **Style Preferences...** in the resulting pull-down menu.
- The resulting dialogue box has two tabs, *Protein* and *Packing*.

- Select the *Packing* tab to specify the style/colour settings for any part of the structure that was generated by application of crystallographic symmetry (i.e. for portions of protein that are in nearby regions of the crystal-packing environment). There may not be any of these.
- Use the *Protein* tab to specify settings for everything else.
- The dialogue box may be used to set default styles and colouring schemes for the various parts of a protein structure. The **Custom Carbon** option allows a colour to be specified for carbon atoms via the colour palette (see Using the Select Color Dialogue Box, page 34), leaving the colours of remaining atoms unchanged.
- Preferences will be retained between sessions.

7.3 Altering Structure Display Properties Globally

7.3.1 Altering the Display Style Globally

- To set all atoms to a new display style (wireframe, stick, ball and stick, or spacefill), you must first ensure that no atoms are selected (see Identifying Selected Atoms, page 26). You can then use any of the following methods:
 - Set the required style in the tool-bar *Style* box, located near the top of the main Hermes window, e.g.



- Right-click in the display-area background, pick **Styles** from the pull-down menu, and select the required style (**Wireframe**, **Capped Sticks**, **Ball and Stick**, **Spacefill**).
- Hit **Display** in the top-level menu, **Styles** in the resulting pull-down menu, then pick the required display style.

7.3.2 Altering the Colouring Scheme Globally

- To colour all atoms in a certain way (e.g. by element type, or in a particular colour), you must first ensure that no atoms are selected (see Identifying Selected Atoms, page 26).
- Then, right-click in the display-area background, pick **Colours** from the pull-down menu, and select the required option.
- The **Custom Carbon** option allows a colour to be specified for carbon atoms via the colour palette (see Using the Select Color Dialogue Box, page 34), leaving the colours of remaining atoms unchanged.
- Colours can also be set for individual atoms and molecules (see Setting Display Properties for Particular Atom(s) or Bond(s), page 32).

7.3.3 Switching Hydrogens On and Off Globally

- The display of hydrogen atoms (if there are any) can be turned on and off by use of the **Show H** and **Hide H** buttons in the Toolbar.
- Alternatively, pick **Display** from the top-level menu, followed by **Show/Hide** and then **Hide hydrogens** in the subsequent pull-down menus.

7.3.4 Switching Unknown Atoms On and Off Globally

- The display of unknown atoms (these may be lone pairs, dummy atoms etc) can be turned on and off using the **Show Unk** and **Hide Unk** buttons in the Toolbar.
- Alternatively, pick **Display** from the top-level menu followed by **Show/Hide** and then **Show Unknown Atoms** or **Hide Unknown Atoms** in the subsequent pull-down menus.

7.3.5 Displaying or Undisplaying Bond Types Globally

- To switch the display of bond-types on or off, use one of the following methods:
 - Right-click in the display-area background and select **Styles** followed by **Display Bond Types** from the resulting pull-down menu.
 - Hit **Display** in the top-level menu, **Styles** in the resulting pull-down menu, then click on **Display Bond Types** from the resulting pull-down menu.
- To identify aromatic rings by displaying a circle within each ring, use one of the following methods:
 - Right-click anywhere in the display-area background and select **Styles** followed by **Display Aromatic Rings** from the resulting menu.
 - Hit **Display** in the top-level menu, **Styles** in the resulting pull-down menu, then click on **Display Aromatic Rings** from the resulting pull-down menu.

7.4 Setting Display Properties for Particular Atom(s) or Bond(s)

There are several methods for setting the display properties of a particular atom, bond, molecule, or set of atoms:

- Select the atoms whose display properties you wish to change (see Ways of Selecting Atoms and Molecules, page 23). Right-click in the display-area background, select the appropriate option from the pull-down menu (**Styles**, **Colours**, **Show/Hide**) and then select the desired display-property setting from the next menu.
- Select the atoms whose display properties you wish to change (see Ways of Selecting Atoms and Molecules, page 23). Hit **Display** in the top-level menu, select the appropriate option from the pull-down menu (**Styles**, **Show/Hide**) then select the desired display-property setting from the next menu.
- Right-click on an individual atom or bond in the 3D display area, select the appropriate option

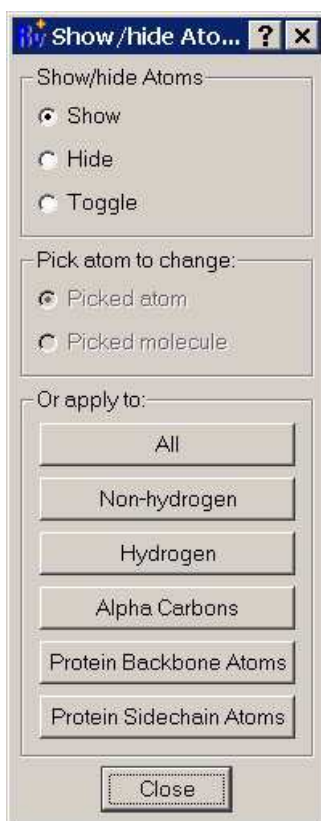
from the pull-down menu (**Styles, Colours, Show/Hide**) and then select the desired display-property setting from the next menu. The chosen setting will be applied to the atom or bond on which you clicked.

- Right-click on an item in the *Protein Explorer* window, select the appropriate option from the pull-down menu (**Styles, Colours**) and then select the desired display-property setting from the next menu. The chosen setting will be applied to all atoms/bond in the chosen item.
- With no atoms selected, hit **Display** in the top-level menu, select **Show/Hide** and then hit **Atoms...** This opens a dialogue box which you can use to (a) specify the desired settings, and then (b) select the atoms or molecule to which those settings are to be applied (see Using the Show/hide Dialogue Box, page 33).

7.5 Using Display-Property Dialogue Boxes

7.5.1 Using the Show/hide Dialogue Box

- You may be presented with the following dialogue box when setting atom display properties:



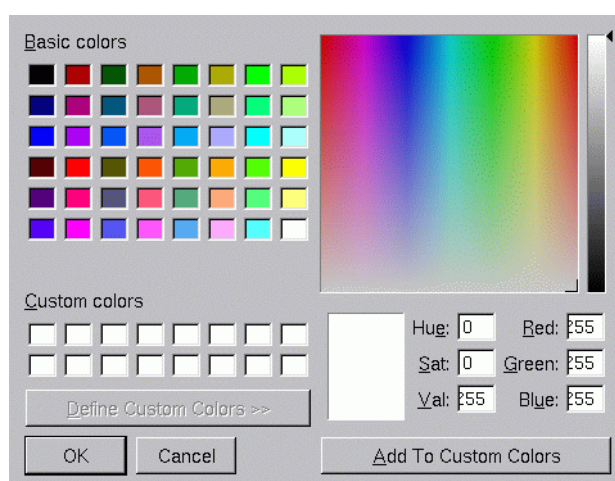
- If there were atoms selected before the dialogue box was opened (see Identifying Selected Atoms, page 26), only the top part of the box will be shown and there will be **Cancel** and **OK** buttons at the bottom instead of **Close**.
- Select the desired display-property setting by switching on the appropriate button in the top area

of the box.

- If atoms were selected before the dialogue box was opened, all you then have to do is hit **OK**. This closes the dialogue box and applies the setting you have chosen to the selected atoms.
- If no atoms were selected before the dialogue box was opened, use the buttons in the lower half of the box to apply the chosen display-property setting to:
 - **Picked atom(s)** - click on the required atoms.
 - **Picked molecule(s)** - click on one atom of the required molecule(s).
 - Particular categories of atoms, e.g. backbone atoms.
- Hit **Close** to close down the dialogue box.

7.5.2 Using the Select Color Dialogue Box

- You may need to use the *Select Color* dialogue box when choosing a colour for an atom, molecule or other object:



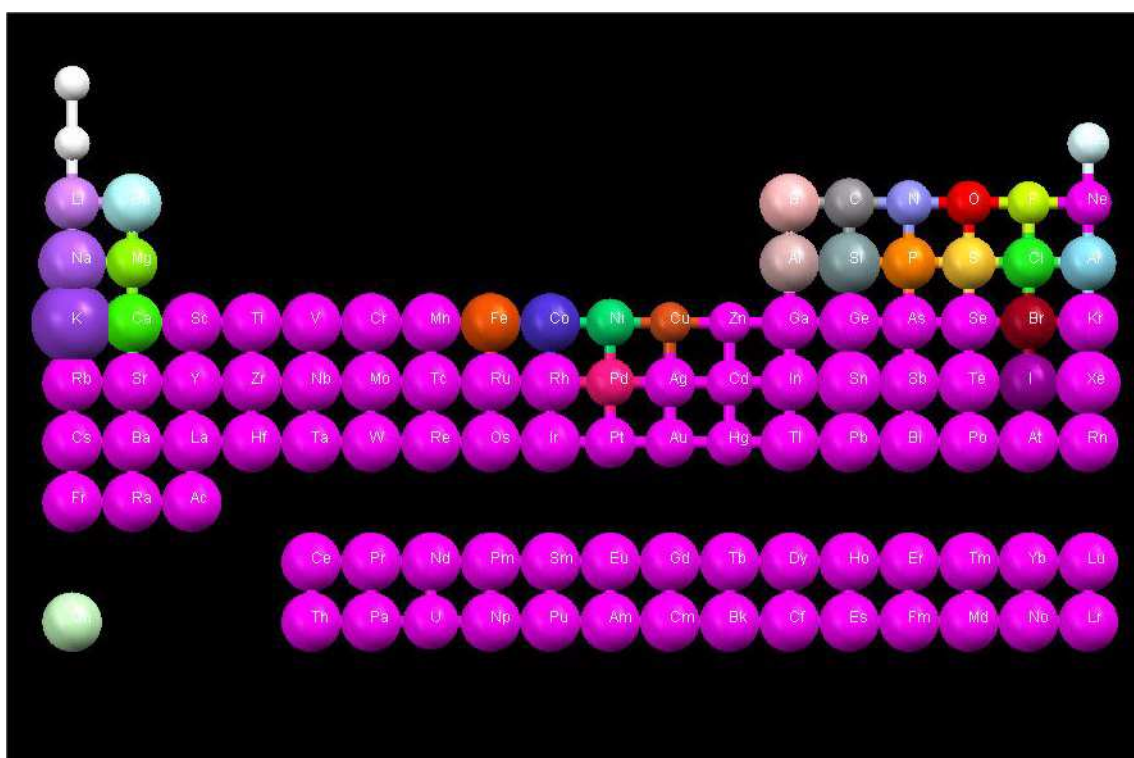
- There are several ways to select a colour:
 - Click on one of the *Basic colors*.
 - Click on the *colour palette*.
 - Click on one of the *Custom colors* (only possible if you have defined some earlier in the Hermes session).
 - Click in the vertical *lightness slider*.
 - Type your required *Red*, *Green* and *Blue* values.
 - Type your required *Hue*, *Saturation* and *Lightness* (i.e. *Val.*) values. The *Hue* defines the colour, e.g. 0 is red, 120 is green, and 240 is blue. The *Saturation* defines the amount of colour; 255 is fully saturated, whereas a colour with a saturation of zero is always white. The *Lightness* can also be controlled; a *Lightness* of zero always gives black.
- Once you have chosen the required colour, hit **OK** to continue.

- To create a customised palette, you can choose a colour and then hit **Add To Custom Colors**. This adds the current colour to the *Custom colors* area for use later in the same session.

7.6 Element and Bond-Type Display Conventions

7.6.1 Standard Element Colours

- When atoms are coloured by element (see *Altering the Colouring Scheme Globally*, page 31), each atom is assigned a colour depending on its element type.
- Colours used for the common elements are:



- Atoms of unknown element type are coloured pale green (bottom left of the view above).
- All other elements are coloured purple.
- There is no way to modify the colour used for a particular element, except for carbon, which may be assigned a customised colour globally (see *Altering the Colouring Scheme Globally*, page 31), for particular parts of the protein (see *Setting Default Style and Colour Preferences for Different Parts of the Protein*, page 30) or for individual atoms or sets of atoms.

7.6.2 Bond-Type Display Conventions

- Single, double, triple and quadruple bonds are indicated by one, two, three and four lines, respectively.
- A circle within a ring indicates that all the ring bonds have the aromatic bond type.

- If some, but not all, bonds in a ring have the aromatic type (which is not chemically reasonable but can happen if a molecule file has been edited by hand), no circle is displayed in the centre of the ring but each aromatic bond is indicated by a small red circle going through the bond.
- Pi bonds to metals are indicated by single broken lines.
- Delocalised bonds are indicated by double broken lines.
- The display of bond types and aromatic rings can be toggled on and off (see [Displaying or Undisplaying Bond Types Globally](#), page 32).

8 Labelling Atoms

8.1 Types of Labels

- Atoms may be labelled by:
 - Their atom name, e.g. *CA* or *NZ*.
 - Their element type, e.g. *S*.
 - The atom type e.g. *C.ar*.
 - The number of the residue to which they belong, e.g. *SER11*.
 - The sequential number of the atom in the protein file, the Atom File Index.
 - The identifier of the protein structure to which they belong, e.g. *IDX4*.

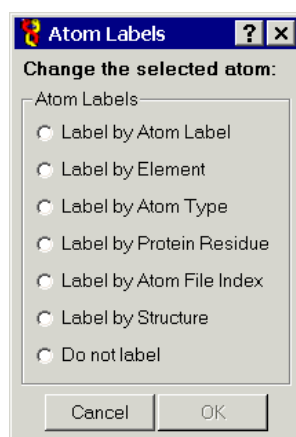
8.2 Turning All Atom Labels On or Off

- To switch all atom labels on or off, you must first ensure that no atoms are selected (see Identifying Selected Atoms, page 26). You can then use any of the following methods:
 - Right-click in the display-area background, pick **Labels** from the pull-down menu, and select the required option (**Label by Atom Label**, **Label by Element**, etc.).
 - Hit **Display** in the top-level menu, **Labels** in the resulting pull-down menu, then pick the required option (**Label all**, **Clear all labels**).
- Atom labels can also be turned on and off individually (see Labelling Individual Atoms and Sets of Atoms, page 37). The positions of atom labels can be optimised (see Optimising Label Positions, page 39).

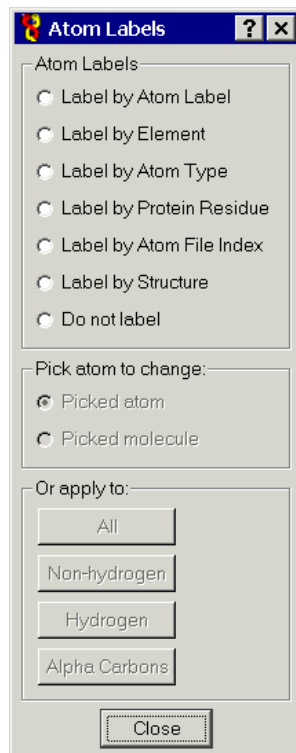
8.3 Labelling Individual Atoms and Sets of Atoms

There are several methods for labelling a particular atom, bond, molecule, or set of atoms:

- Select the atoms whose labels you wish to set (see Ways of Selecting Atoms and Molecules, page 23). Right-click in the display-area background, pick **Labels** from the pull-down menu, and select the required option (**Label by Atom Label**, **Label by Element**, etc.).
- Select the atoms whose labels you wish to set (see Ways of Selecting Atoms and Molecules, page 23). Hit **Display** in the top-level menu, then pick **Labels** followed by **Atoms...** in the resulting pull-down menus. Select the required option from the dialogue box:



- Right-click on an individual atom or bond, or on an item in the *Protein Explorer* window, select **Labels** from the pull-down menu and then select the desired option from the next menu. The chosen setting will be applied to the atom or bond on which you clicked, or all atoms/bond in the chosen *Protein Explorer* item.
- With no atoms selected, hit **Display** in the top-level menu, then **Labels** followed by **Atoms...** in the resulting pull-down menus. This opens a dialogue box which you can use to (a) specify the desired label types, and then (b) select the atoms or molecule or types of atoms to be labelled.



8.4 Customising the Appearance of Labels

8.4.1 Setting the Colours of Labels

To set the colour of all labels, use one of the following methods:

- Hit **Display** in the top-level menu, then **Display Options...** in the menu. In the resulting *Display Options* dialogue select the *Label* tab, hit the coloured panel next to **Label Color** and select from the resulting colour palette (see Using the Select Color Dialogue Box, page 34).
- Hit **Display** in the top-level menu, **Labels** in the resulting pull-down menu, then **Label colour...** and select from the resulting colour palette (see Using the Select Color Dialogue Box, page 34).
- Right-click in the display-area background, pick **Labels** from the pull-down menu, select **Label colour...** and choose from the resulting colour palette (see Using the Select Color Dialogue Box, page 34).

8.4.2 Setting Label Sizes

To set the size of all labels use one of the following methods:

- Hit **Display** in the top-level menu, then **Display Options...** in the pull-down menu. In the resulting *Display Options* dialogue select the *Label* tab and enter a new **Label Size**.
- Hit **Display** in the top-level menu, **Labels** in the resulting pull-down menu, then **Label size...** and enter a new label size in the resulting dialogue.
- Right-click in the display-area background, pick **Labels** from the pull-down menu, select **Label size...** and enter a new label size in the resulting dialogue.

8.4.3 Optimising Label Positions

- The Hermes visualiser can try to optimise the positions of atom labels to minimise overlap with other graphics objects such as atoms. The aim is to produce a clearer display, e.g., for publication. However, because proteins have so many atoms, the optimisation procedure will often be unable to produce a good result.
- To optimise the positions of atom labels, hit **Display** in the top-level menu, then **Display Options...** in the pull-down menu. In the resulting *Display Options* dialogue select the *Labels* tab and enable the *Optimise Atom Label Positions* check-box.

9 Calculating and Displaying Ribbons, Tubes and Other Graphics Objects

9.1 Creating Centroids

The centroid of a set of atoms can be created in two ways:

Method 1:

- Choose **Calculate** from the top-level menu and **Centroids...** from the pull-down menu. Hit **New Centroid...** in the resulting dialogue box.
- This opens the *Centroid Properties* dialogue box, which allows you to create the least-squares centroid of:
 - A set of two or more atoms: switch on the **Picked atoms** button, click on the required atoms in the display, then hit **OK**.
 - A molecule or a set of molecules: switch on the **Picked molecule** button, click on one atom in each of the required molecules, then hit **OK**.
 - **All atoms** in the display, all **Non-hydrogen** atoms in the display, or all **Hydrogen atoms** in the display.
- You can use the options in the bottom part of the *Centroid Properties* dialogue box to set display properties such as label and colour (see Setting the Display Properties of Centroids, page 41).

Method 2:

- Select two or more atoms (see Ways of Selecting Atoms and Molecules, page 23).
- Right-click anywhere in the display-area background, select **New** from the pull-down menu and then **Centroid** in the next menu.
- The new centroid will appear. Its display properties (colour, label, etc.) can be edited if required (see Setting the Display Properties of Centroids, page 41).

9.2 Setting the Display Properties of Centroids

Centroids have several display properties, viz. colour; transparency (i.e. whether you can see through them or not); label (i.e. what the centroid is called); whether the label is displayed or not; whether the centroid itself is displayed or hidden. These properties may be controlled via the *Graphics Objects Explorer* window (see Controlling the Display of Centroids, Ribbons and Contoured Surfaces: the Graphics Objects Explorer, page 48) or in one of the following ways:

Method 1:

- Choose **Calculate** from the top-level menu and **Centroids...** from the next menu.
- The resulting dialogue box lists the current set of centroids, e.g.



- Left-click on the required centroid and hit **Edit...**
- Use the resulting *Properties* dialogue box to set the desired display properties. Clicking on the coloured panel will open the *Select Color* dialogue box which can be used for customising colours (see Using the Select Color Dialogue Box, page 34).
- In the example below, the user has set *Show* on (i.e. the object will be displayed, not hidden), the label on (i.e. a label *MY CENTROID* will be displayed), the transparency off (i.e. the centroid will be opaque) and the colour to green:



Method 2:

- Right-click on the centroid in the display area, select **Colours**, **Labels** or **Show/Hide** as required, then select the desired option.
- For example:
 - Selecting **Label** followed by **Edit label...** will allow you to type in a new name for the object.
 - Selecting **Colours** followed by **Blue** will turn the object blue.
 - Selecting **Show/Hide** followed by **Hide** will undisplay the object.
 - Selecting **Colours** followed by **Colour...** will open the *Select Color* dialogue box which can be

used for customising colours (see Using the Select Color Dialogue Box, page 34).

Method 3:

- Choose **Display** from the top-level menu, then select **Labels**, **Colours** or **Show/hide** in the pull-down menu, depending on which property you wish to alter.
- Select **Centroids...** from the next menu.
- The resulting dialogue box lists the current set of centroids:
 - For colours, left-click on the item in the list that you wish to change. This will open the *Select Color* dialogue box which you can use to choose the required colour (see Using the Select Color Dialogue Box, page 34).
 - For labels and show/hide, set the *Action* buttons to the required setting and then left-click on the item in the list that you wish to change.
- Alternatively, right-clicking on an item in the list will produce the menus described above under *Method 2*.

Note: Changes made to the colour and/or size of centroid labels will be applied to all labels in the display (i.e. all atom and centroid labels).

9.3 Deleting Centroids

Centroids can be deleted (i.e. completely removed, not hidden) in several ways:

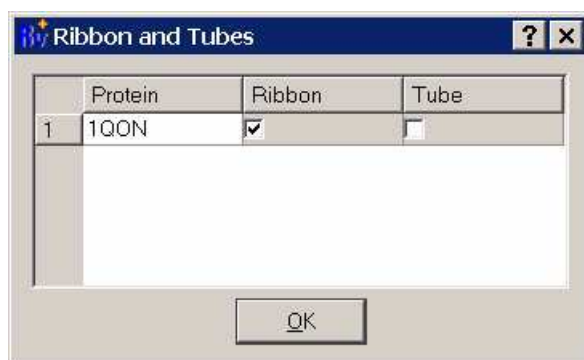
- Via the *Graphics Objects Explorer* dialogue (see Controlling the Display of Centroids, Ribbons and Contoured Surfaces: the Graphics Objects Explorer, page 48).
- Right-click on a centroid in the display. In the resulting pull-down menu, select either **Delete Object** to delete the centroid you have clicked on, or **Clear All Objects** to delete all auxiliary graphics objects such as centroids and ribbons.
- Choose **Calculate** from the top-level menu and select **Centroids...** as appropriate. In the resulting dialogue box:
 - left-click on an item in the list box and hit **Delete**;
 - or right-click on an item in the list and select either **Delete Object** to delete the centroid you have clicked on, or **Clear All Objects** to delete all centroids and other auxiliary graphics objects.

9.4 Creating Ribbon and Tube Displays

Ribbons or tubes can be displayed along protein chains to make secondary structure more obvious:

- Select **Display** from the top-level menu followed by **Ribbons & Tubes...** in the resulting pull-down menu.
- The resulting dialogue lists all proteins currently loaded. Switching on one of the check boxes opposite a given protein identifier will show the ribbon or tube display for that protein, e.g. the

following will show the ribbon display for 1QON:



- It is not possible to show a ribbon and a tube simultaneously for a given protein.
- It is not possible to show ribbons or tubes for some but not all chains in a multi-chain protein.

9.5 Setting the Display Properties of Ribbons and Tubes

The colour and transparency of ribbons and tubes may be changed as follows:

- Via the *Graphics Objects Explorer* dialogue (see Controlling the Display of Centroids, Ribbons and Contoured Surfaces: the Graphics Objects Explorer, page 48).
- Right-click on the ribbon or tube in the display area, select **Colours**, then select the desired option.
- For example:
 - Selecting **Colours** followed by **Blue** will turn the ribbon or tube blue.
 - Selecting **Colours** followed by **Colour...** will open the *Select Color* dialogue box which can be used for customising colours (see Using the Select Color Dialogue Box, page 34).
 - Selecting **Transparent** will make it possible to see through the ribbon or tube; **Opaque** will make the object non-transparent.

9.6 Deleting Ribbons and Tubes

Ribbons and tubes may be deleted by:

- Right-clicking on the **Ribbon Section** submenu option in the *Graphics Objects Explorer* dialogue and picking **Delete** from the resultant list of options (see Controlling the Display of Centroids, Ribbons and Contoured Surfaces: the Graphics Objects Explorer, page 48).
- Right-clicking on the object and selecting **Delete Object** in the resulting pull-down menu. **Clear All Objects** in this menu will delete all ribbons and tubes, together with any other auxiliary objects, such as centroids.
- Right-clicking on the object and selecting **Show/Hide** in the resulting pull-down menu, followed by **Hide** in the next pull-down.
- Switching off the appropriate check box in the *Ribbons & Tubes* dialogue box (if the dialogue is

not already open, it can be accessed by hitting **Display** in the top-level menu followed by **Ribbons & Tubes** in the pull-down menu).

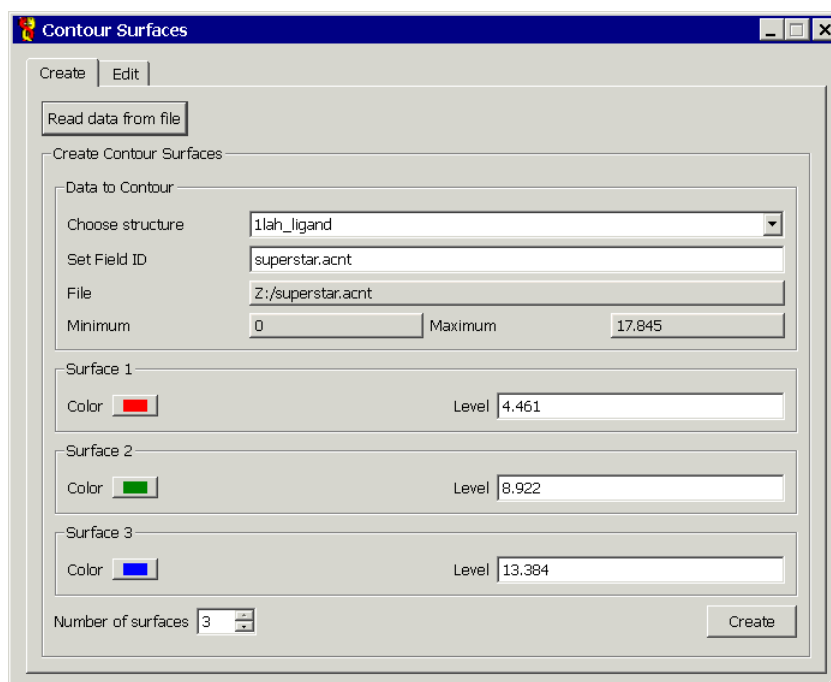
9.7 Visualising Contoured Surfaces

Contoured surfaces can either be generated for a loaded protein using SuperStar (see SuperStar documentation for further details) or by loading a contour surface file generated using another package in addition to the corresponding structure (see below).

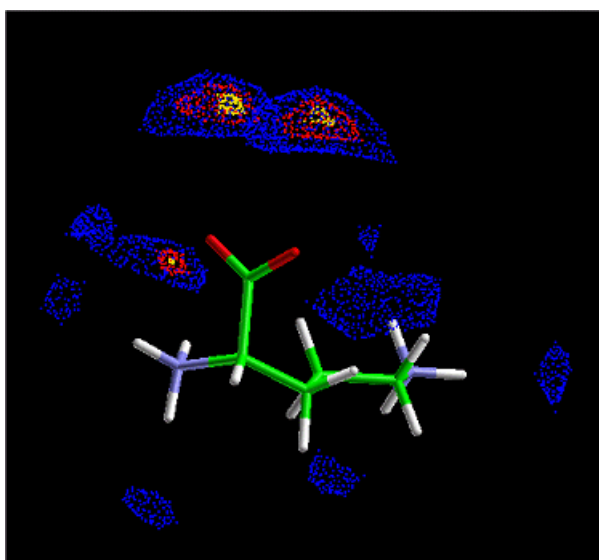
- Contoured surface files generated for small molecules or proteins can be read into the Hermes visualiser (along with the corresponding structure) and visualised/edited.
- Contour files must be in ASCII contour file (.acnt) format.
- These files are generally output from packages such as IsoStar (http://www.ccdc.cam.ac.uk/products/csd_system/isostar/) or SuperStar (http://www.ccdc.cam.ac.uk/products/life_sciences/superstar/).
- The structure from which the contour file has been generated can be read into the visualiser using **File** then **Open**. Navigate to the directory containing the structure file, select the file and click **OK**.

Note: if more than one file has been loaded into Hermes, you will have to select the appropriate file from the *Choose Structure* pull down.

- To read in the contour file, select **Display** and then **Contour Surfaces**. Click on the **Read data from file** button and load the .acnt file in the same way as the structure file above. The window will update to reflect the file that has been read in.



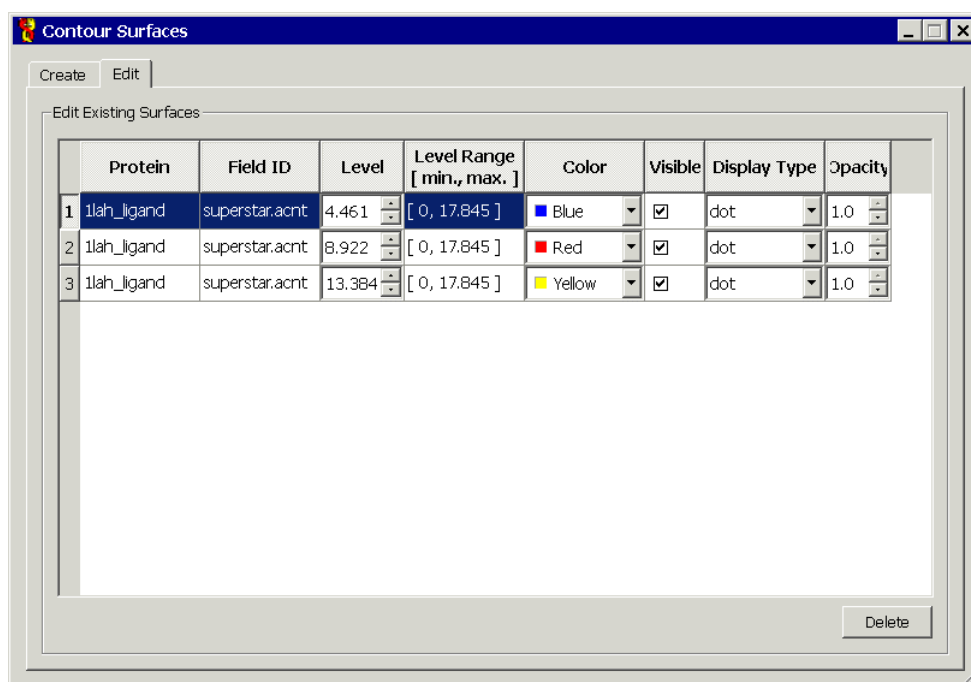
- Initially no contours are displayed. You can edit the following criteria before creating the contour map:
 - The contour level range (maximum and minimum values can be specified). Note that Hermes provides an optimal range by default that has been extracted from the contour file.
 - The number of surfaces displayed: controlled using the **Number of surfaces** option at the bottom of the interface.
 - The surface colours: click on the **Color** button and pick from the palette or define your own colour (see Using the Select Color Dialogue Box, page 34). Hit **OK** to return to the *Contour Surfaces* window.
 - The contour levels for the different contours: these can be scaled based on the contour level for *Surface 1*. The greater the number, the greater the propensity for a contact occurring at this location.
- Hit the **Create** button to view the contour surfaces.



- Contours can be edited once they have been generated (see Setting the Display Properties for Contoured Surfaces, page 46).

9.8 Setting the Display Properties for Contoured Surfaces

- Once contoured surfaces have been generated, they can either be edited via the *Edit* tab on the *Contour Surfaces* window:



or via the *Graphics Objects Explorer* dialogue (see Controlling the Display of Centroids, Ribbons and Contoured Surfaces: the Graphics Objects Explorer, page 48).

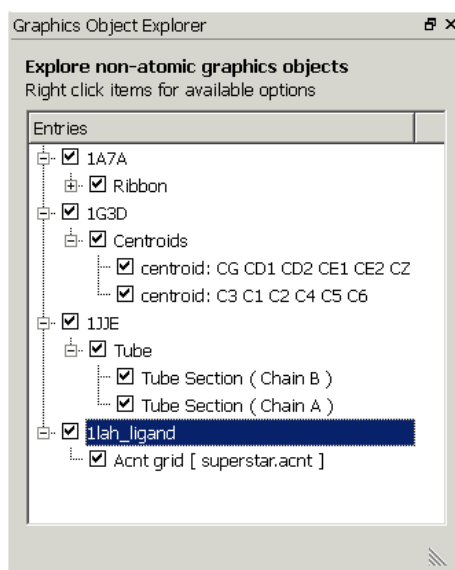
- Parameters that can be modified include:
 - Contour level: increase or decrease using the arrow keys in the **Level** menu, or type the new value (you'll need to hit the **Return** key to impose the changes).
 - Contour colour: hit the coloured entry (i.e. blue, red or yellow shown above) in the table, select the new colour then click **OK**.
 - Display Type: use the pull-down menu to select from *dot*, *dot (less)*, *dot (more)*, *edge*, *node* or *triangle* display types.
 - Opacity: the default value is 1. Enter a number between 0 and 1 and hit the **Return** key, or use the arrow keys to select the new opacity value.

9.9 Deleting Contoured Surfaces

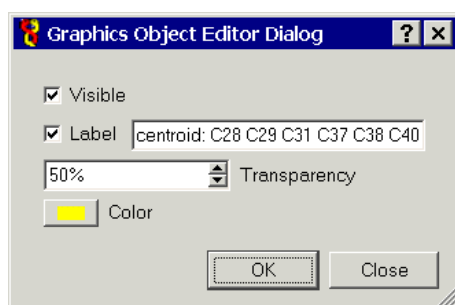
- Contours can be deleted by selecting the contour row and hitting the **Delete** button at the bottom of the *Contour Surfaces* window.
- This can also be done via the *Graphics Objects Explorer* Window by right-clicking on the **Propensity Map** entry in the tree view, hitting **Edit** then selecting the contour to be deleted from the resultant window. Alternatively all contours can (see Controlling the Display of Centroids, Ribbons and Contoured Surfaces: the Graphics Objects Explorer, page 48).

9.10 Controlling the Display of Centroids, Ribbons and Contoured Surfaces: the Graphics Objects Explorer

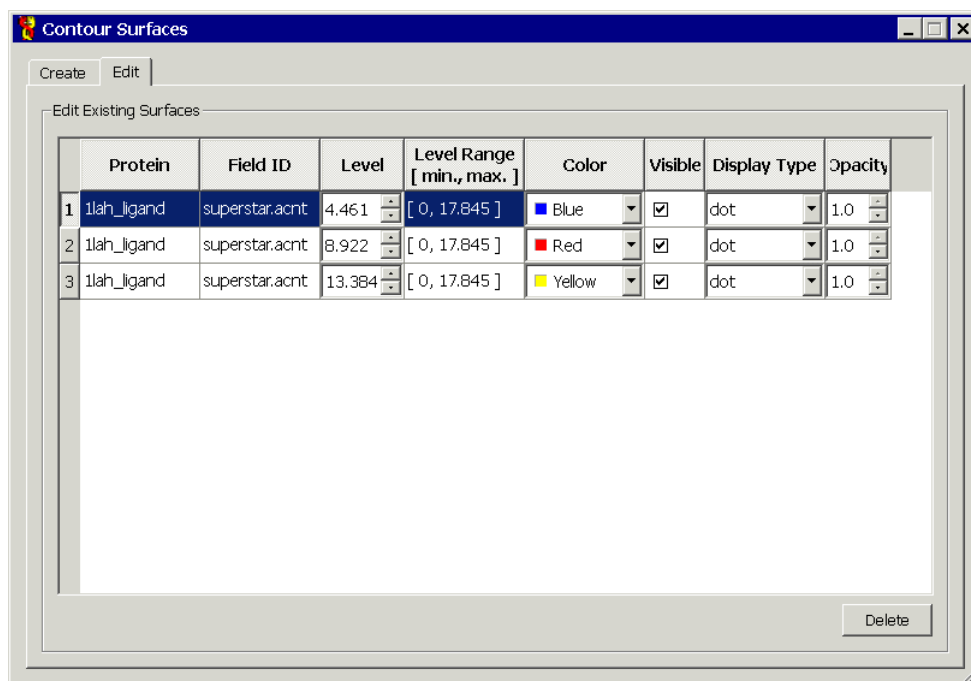
- The Graphics Objects Explorer can be launched either using the Toolbar **Graphics Objects** button or via the **Display, Graphics Objects Explorer...** main menu option.
- Any non-atomic graphics objects such as centroids (see Creating Centroids, page 41), ribbons or tubes (see Creating Ribbon and Tube Displays, page 43) and contour plots (see Visualising Contoured Surfaces, page 45) that have been generated will be listed in this dialogue window.



- The way in which graphics objects are displayed (i.e. colours, settings, or even if the object is displayed at all) is controlled using this dialogue.
- These settings can be modified by right-clicking on the object to be edited or deleted.
Note: the appearance of the resultant pull-down menu will vary depending on how far down the tree you are, e.g. from the dialogue above, if you right-click on **Tube** (underneath **1JJJ**), you will only be able to modify the *Colours* of all tubes associated with the particular protein (i.e. the **Edit** and **Delete** menu options will be greyed out). To delete the object, or edit it more comprehensively, it is necessary to expand the tree further (e.g. using the + to the left of *Tube*), then right-click on the text underneath *Tube* on the tree (for this example, the text would be of the format *Ribbon Section (Chain A)*)).
- To modify settings for a particular graphics object:
 - Right-click and select **Edit** from the resultant pull-down menu. For centroids, ribbons and tubes you will be able to modify the visibility (whether the object is displayed or not), the label applied to the object, the transparency and the colour.



- For contour plots you will be able to modify propensity values, contour colours, whether or not specific contours are displayed, the display type and the opacity. These settings are described in more detail elsewhere (see Setting the Display Properties for Contoured Surfaces, page 46).



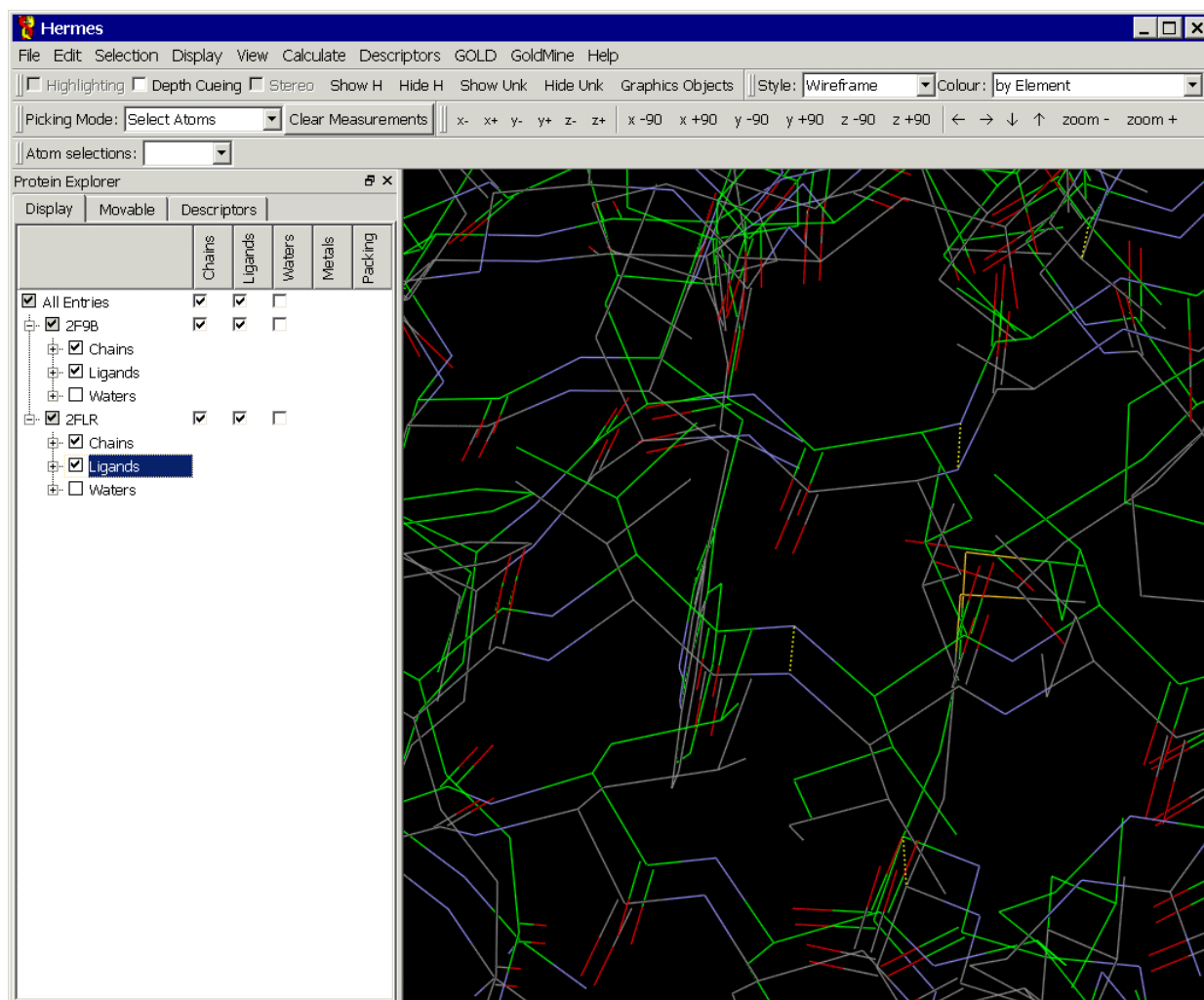
- Click **OK** to close the dialogue window once you have applied any changes.
- To delete a graphics object, right-click on the chosen object and select **Delete** from the pull-down menu. This will remove the object from the 3D view and also from the *Graphics Objects Explorer* window.
- The *Graphics Objects Explorer* window is dockable which means it can be placed in various positions in the main Hermes window, or kept as an entirely separate window. To do this, move the mouse cursor into the top left of the window so that the four-arrow icon appears then drag the window with the mouse, keeping the left-button depressed, and put it down where you want by releasing the mouse button.

10 Overlaying or Superimposing Proteins

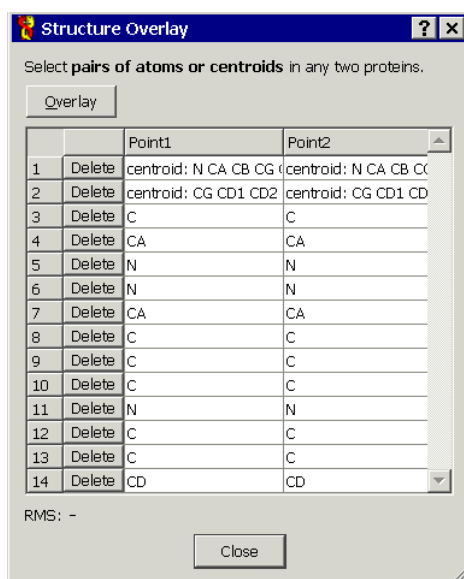
- Structures can be overlayed one of two ways within Hermes:
 - By least squares overlay by selection of pairs of atoms (see Overlaying Structures, page 51).
 - By sequence alignment (see Superimposing Protein Structures, page 52).

10.1 Overlaying Structures

- To perform least-squares overlay of pairs of structures, read the two protein structures into Hermes so that they are visible in the 3D display. Select **Calculate** from the top-level menu followed by **Structure Overlay...** from the resulting pull-down menu. This will open the *Structure Overlay* dialogue.
- In the Hermes 3D view select pairs of atoms from the two proteins you wish to overlay (these pairs of atoms will be used to calculate the least-squares overlay). Broken yellow lines will be displayed between each selected pair of atoms:



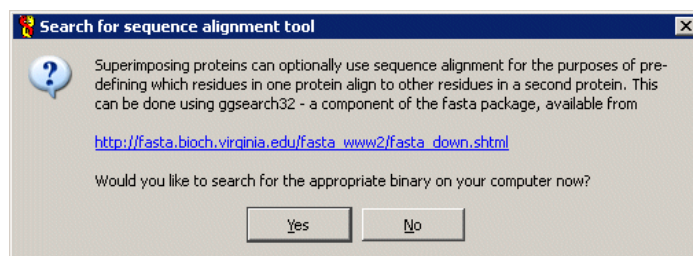
- As pairs of atoms or centroids are selected they will be listed in the *Structure Overlay* dialogue:



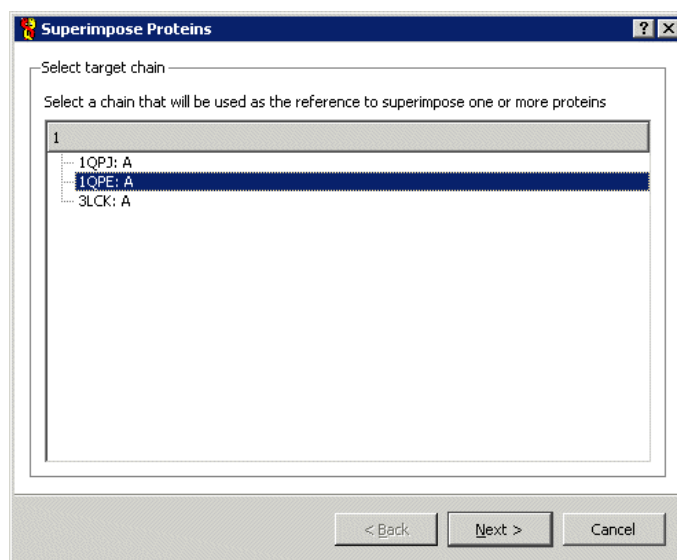
- Once the desired number of atom/centroid pairs are selected, hit the **Overlay** button.
- Overlay is performed by moving the second crystal structure relative to the first crystal structure. During overlay crystal structures will be kept rigid, e.g. it is not possible to rotate torsions to find better fits.
- The root-mean-square (RMS) of the pairs of atoms used for the overlay is displayed in the *Structure Overlay* dialogue.
- To refine the orientation, you can continue to select further pairs of atoms. These will be added to the *Structure Overlay* dialogue. To remove a pair of atoms hit the **Delete** button next to the pair you wish to remove. Hit the **Overlay** button in order to re-calculate the overlay using the current selection.

10.2 Superimposing Protein Structures

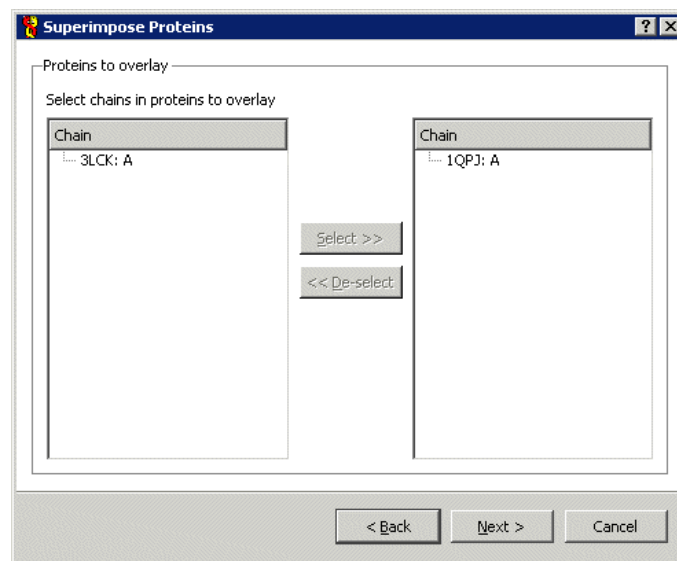
- Superimposing protein structures in Hermes can be done using one of two methods, both of which use sequence alignment: FASTA or the Needleman-Wunsch algorithm. Both FASTA and the Needleman-Wunsch algorithm do the same thing: i.e. they generate global sequence alignments which give a pair-wise matching of one residue to another which can then be used for overlay.
- To overlay two or more protein structures, read the protein structures into Hermes so that they are visible in the 3D display. Select **Calculate** from the top-level menu followed by **Superimpose Proteins ...** from the resulting pull-down menu. This will open a *Search for sequence alignment tool* window.



- If you have FASTA installed on your PC, click on **Yes** and use the file browser to locate the appropriate binary on your PC. If you do not have FASTA installed it can be downloaded from the URL provided: http://fasta.bioch.virginia.edu/fasta_www2/fasta_down.shtml.
- If you wish to use the default Needleman-Wunsch algorithm, click **No**.
- Once the method for superimposing is selected a *Superimpose Proteins* window will guide you through the superposition process.
- You must first select the protein that will be used as the reference structure.



- After the reference protein has been selected, select the protein(s) that you wish to overlay with the reference.



- Before superposition can be done, a residue filter must be selected:
 - *Use whole protein*: use the sequences in the reference protein as a basis for the overlay.
 - *Limit to target residues in selection*: select a previously saved selection (see Making Complex Selections: Defining Subsets of Protein Atoms, page 24) to be used as the basis for the overlay.
 - *Select residues manually*: select pairs of residues from use the scrolling list of sequences as a basis for the overlay.
- Once you have selected the residue filter, click Next then Transform Structures to perform the overlay. If you wish to change any of the settings prior to superposition use the **Back** button.

11 Measuring Geometrical Parameters

11.1 Measuring Distances, Angles and Torsions

- Measurement mode can be set in two ways:
 - By setting the *Picking Mode* in the tool bar (near the top of the main Hermes window) to the required parameter type, viz. **Measure Distance**, **Measure Angle** or **Measure Torsion**.
 - By right-clicking in the display area (but not on an atom, bond or object), selecting **Measure** from the pull-down menu, and then choosing **Measure Distances**, **Measure Angles** or **Measure Torsions**.
- Geometrical measurements (intramolecular or intermolecular) can now be displayed by clicking on e.g., two atoms for a distance, three atoms for an angle or four atoms for a torsion angle.
- Centroids can be used in distance, angle and torsion angle measurements.
- The colour of the dotted line showing a measurement can be altered by:
 - Right-clicking on any part of the dotted line that does not lie on a chemical bond;
 - Selecting **Colours** from the resulting pull-down menu;
 - Picking the desired colour.

11.2 Clearing Measurements

- All geometrical measurements can be removed from the display by:
 - Clicking on **Clear Measurements** in the tool bar near the top of the main Hermes window.
 - Right-clicking in the display-area background and then hitting **Clear Measurements** in the resulting pull-down menu.
 - Right-clicking in the display-area background, selecting **Measure** in the resulting pull-down menu, and then hitting **Clear Measurements** in the next menu.
- An individual measurement can be removed by right-clicking on any part of the green dotted line that does not lie on a chemical bond and selecting **Delete Measurement** from the resulting pull-down menu.

12 Finding Hydrogen Bonds and Other Nonbonded Contacts

12.1 Accessing Nonbonded-Contact Functionality

There are two basic ways in which nonbonded contacts can be found and highlighted by the Hermes visualiser:

- By use of the *Contacts* dialogue box, which provides functionality for finding and displaying hydrogen bonds and clashes (i.e. short nonbonded contacts that are not hydrogen bonds) between:
 - ligand atoms and atoms of the protein;
 - ligands atoms and water molecules;
 - protein atoms and water molecules.Contacts between pairs of protein atoms will not be found.
- By use of the **Select Atoms Within Range** option, which will select all atoms within a user-defined distance of a specified atom or set of atoms.

12.2 Using the *Contacts* Window

12.2.1 Overview

- The *Contacts* window is opened by selecting **View** from the top-level menu followed by **Contacts...** in the resulting pull-down menu (if there is a tick opposite **Contacts...** in this menu, the window is already open).
- The *Contacts* window is *dockable*, which means it can be placed in various positions in the main Hermes window, or kept as an entirely separate window. To do this, move the mouse cursor into the top left of the window so that the four-arrow icon appears.
- Then drag the window with the mouse, keeping the left-button depressed, and put it down where you want by releasing the mouse button.
- The *Contacts* window provides functionality for finding:
 - finding and displaying hydrogen bonds and nonbonded clashes.
 - controlling the definition of what constitutes a hydrogen bond.
 - controlling the definition of what constitutes a nonbonded clash.

12.2.2 Finding Hydrogen Bonds and Nonbonded Clashes

- Open the *Contacts* window (see Using the Contacts Window, page 57).
- If desired, customise the definition of hydrogen bonds (see Changing the Definition of Hydrogen-Bonds, page 58) and nonbonded clashes (see Changing the Definition of Non-bonded Clashes, page 59).
- Switch on the *Ligand Sphere* check box if you want the calculation to be confined to regions

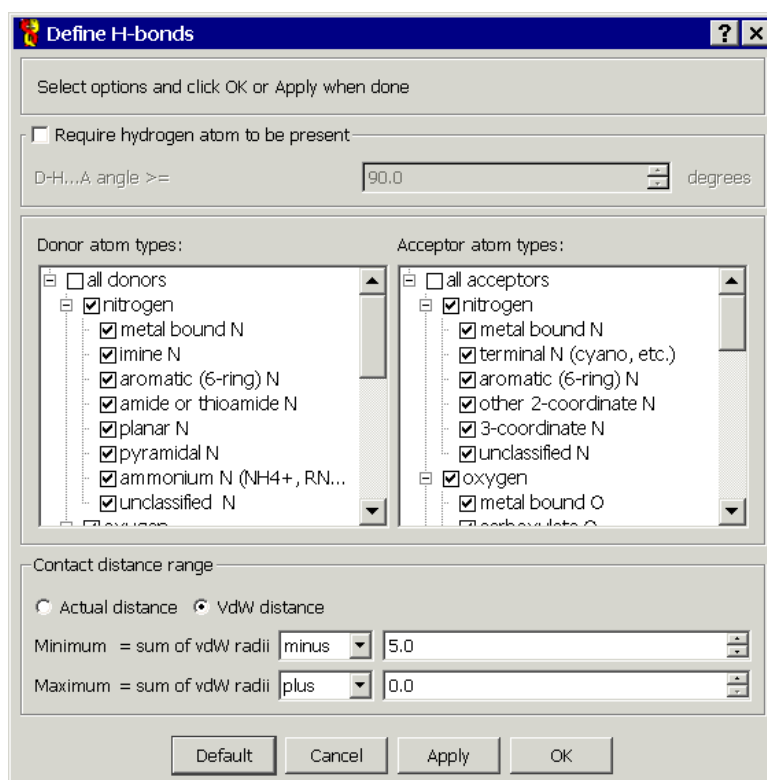
around ligands.

- All currently-loaded proteins will be listed in the *Contacts* window. Switch on the *H-Bonds* check box opposite a particular protein to display, for that protein, hydrogen bonds between:
 - ligand atoms and atoms of the protein.
 - ligands atoms and water molecules.
 - protein atoms and water molecules.Contacts between pairs of protein atoms will not be found.
- Similarly, use the *Short Contacts* check boxes to display short nonbonded contacts that are not hydrogen bonds.

12.2.3 Changing the Definition of Hydrogen-Bonds

- The default definition of a hydrogen bond ($D-H...A$) is a contact meeting the following criteria:
 - The donor (D) must be a nitrogen, oxygen or sulphur atom covalently bound to at least one hydrogen. It is not necessary for the 3D coordinates of the H atom to be known, i.e. H-bonds will be found even if H atoms have not been located in the crystal structure.
 - The acceptor (A) must be a nitrogen, oxygen, sulphur or halogen with at least one available lone pair. Only ionic and metal-bound halogens are considered as acceptors.
 - The $D...A$ distance must be less than the sum of the van der Waals radii of the D and A atoms.
- The criteria used to define what constitutes a hydrogen bond may be altered in the *Contacts* window as follows:
 - Hit the **Define H-Bonds...** button (located near to the top-left of the *Contacts* window).
 - In the resulting *Define H-bonds* dialogue box, specify whether it is necessary for the position of the hydrogen atom involved in the H-bond to be known (i.e. for the H atom to have been located in the crystal-structure analysis). For protein X-ray structures, it will usually be appropriate to allow the H atom to be missing.
 - If the *Require hydrogen atom to be present* button is set on, then a minimum value for the $D-H...A$ angle must be specified (in degrees).
 - Select the required combination of donor and acceptor atoms via the tick-boxes provided in the scrolling windows labelled *Donor atom types* and *Acceptor atom types*. It is not possible to guarantee perfect classification of donors and acceptors if one or more H atoms are missing from the structure, as will normally be the case for protein X-ray structures. C-H is not available as a donor group.
 - Set the required upper and lower distance limits. These may be specified as absolute distances in Angstroms, or as distances relative to the sum of vdW radii of the atoms involved in the H-bond. If the *Require hydrogen atom to be present* button is on, the distance constraint refers to the $H...A$ distance; otherwise it refers to the $D...A$ distance.
 - Hit **Apply** or **OK** to complete the hydrogen-bond definition (the former leaves the dialogue box open, the latter closes it).

- Hit **Default** followed by **Apply** or **OK** if you want to return all the settings to their default values.
- The following settings will find H-bonds between all types of donors and 6-membered ring aromatic nitrogen acceptors, subject to the constraint that the donor-atom...nitrogen distance must be at least 0.1Å shorter than the sum of vdW radii:



12.2.4 Changing the Definition of Non-bonded Clashes

- The default definition of a short contact in the Hermes visualiser is any contact shorter than the sum of the van der Waals radii of the atoms involved, but excluding anything that is classified as a hydrogen bond.
- This definition can be changed in the *Contacts* window as follows:
 - Hit the **Define Short Contacts...** button (located near to the top-right of the *Options Window*).
 - Specify the require distance limit in the resulting dialogue box, e.g. the following setting would define as a clash any non-H-bonding contact whose distance is less than $v+0.2\text{\AA}$, where v is the sum of the van der Waals radii of the atoms involved:



- Hit **Apply** or **OK** to complete the definition (the former leaves the dialogue box open, the latter closes it).
- Hit **Default** followed by **Apply** or **OK** if you want to return all the settings to their default values.

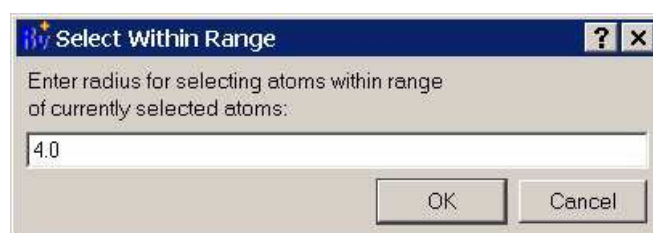
12.2.5 Clearing the Display of Non-bonded Contacts

The display of H-bonds and clashes can be turned off in the same way as for other measurements (see Clearing Measurements, page 55).

12.3 Finding Atoms Within a Specified Distance Range

To identify all atoms within a specified distance of a particular atom or set of atoms:

- Select the atom or set of atoms.
- Pick **Selection** in the top-level menu, **Select** in the pull-down menu, then **Set Selection Radius** in the next menu. Type the required distance range into the resulting entry box and hit **OK**:



- Then pick **Selection** from the top-level, **Select** in the pull-down, and either **Select Atoms Within Range (all proteins)** or **Select Atoms Within Range (selected proteins)**. The former will find and select all atoms within the specified distance of any of the atoms you picked, irrespective of which protein structure they belong to. The latter will only find atoms in proteins that contain one of the atoms you picked.
- Alternatively, the same options can be accessed by right-clicking on an atom in the display area and picking **Selection** followed by **Select** from the resulting pull-down menus.

13 Visualising CSD- and PDB-based Intramolecular Contact Distributions

13.1 Overview

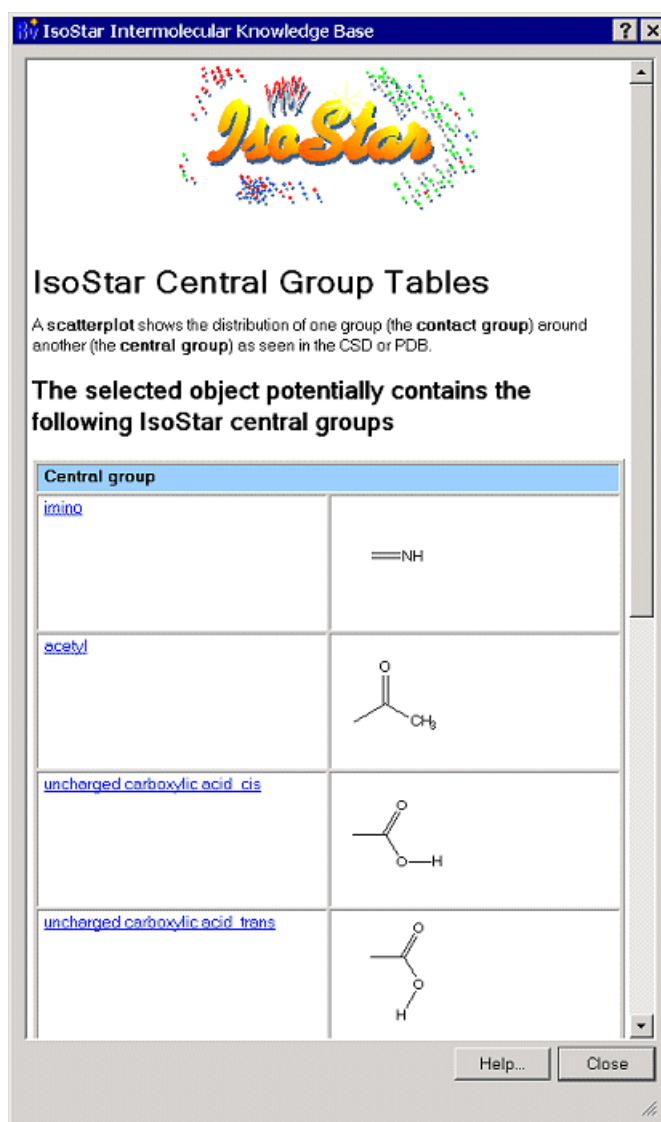
- IsoStar is a library of intramolecular interactions derived from the Cambridge Structural Database (CSD) and provides information on the frequency with which a non-bonded interaction occurs and the preferred orientation of the contact.
- IsoStar *scatterplots* (the distribution of a contact group B around a superimposed central group A) can be used to validate e.g. the geometry of an H-bond or short range interaction between the protein and ligand.
- Scatterplots can be retrieved for either a protein residue (or residues) or the ligand. In each case, the selected molecule or residue will be broken down into constituent functional groups (known as central groups within IsoStar). Once the fragment of interest is selected by the user, a list of available contact groups is presented. The scatterplot is loaded into the client window once the contact group is selected.
- The IsoStar functionality is only available to registered Cambridge Structural Database (CSD) System subscribers. For further information on obtaining the CSD, please see:
http://www.ccdc.cam.ac.uk/contact/obtaining_products/.
- For more information on IsoStar, please visit
http://www.ccdc.cam.ac.uk/products/csd_system/isostar/.

13.2 Setting Up Hermes For Use With IsoStar

- The location of the server must be specified so that the scatterplots can be successfully retrieved.
- Go to the main menu option **Display**, then select **IsoStar Setup** and insert your server location in the resultant *IsoStar Settings* dialogue, e.g. http://computer_name/html/isostar.html.

13.3 Selecting the Central Group

- Select the residue or ligand atom (or atoms) of interest. Please note that it may first be necessary to add H atoms to the ligand (see Molecule Standardisation, page 65).
- Once an atom (or atoms) is selected, either go to the **Calculate** pull down menu and select **IsoStar Intermolecular Contact Database**, or right-click on one of the selected atoms and pick the **IsoStar Intermolecular Contact Database** option from the resultant pull-down menu.
- A pop-up window will appear, detailing the central groups that were found in the residue or molecule that was selected.



- Pick the preferred central group by clicking on the *Central group* hyperlink, e.g. *uncharged carboxylic acid* above. Picking a central group will take you to the a page detailing the list of available contact groups.

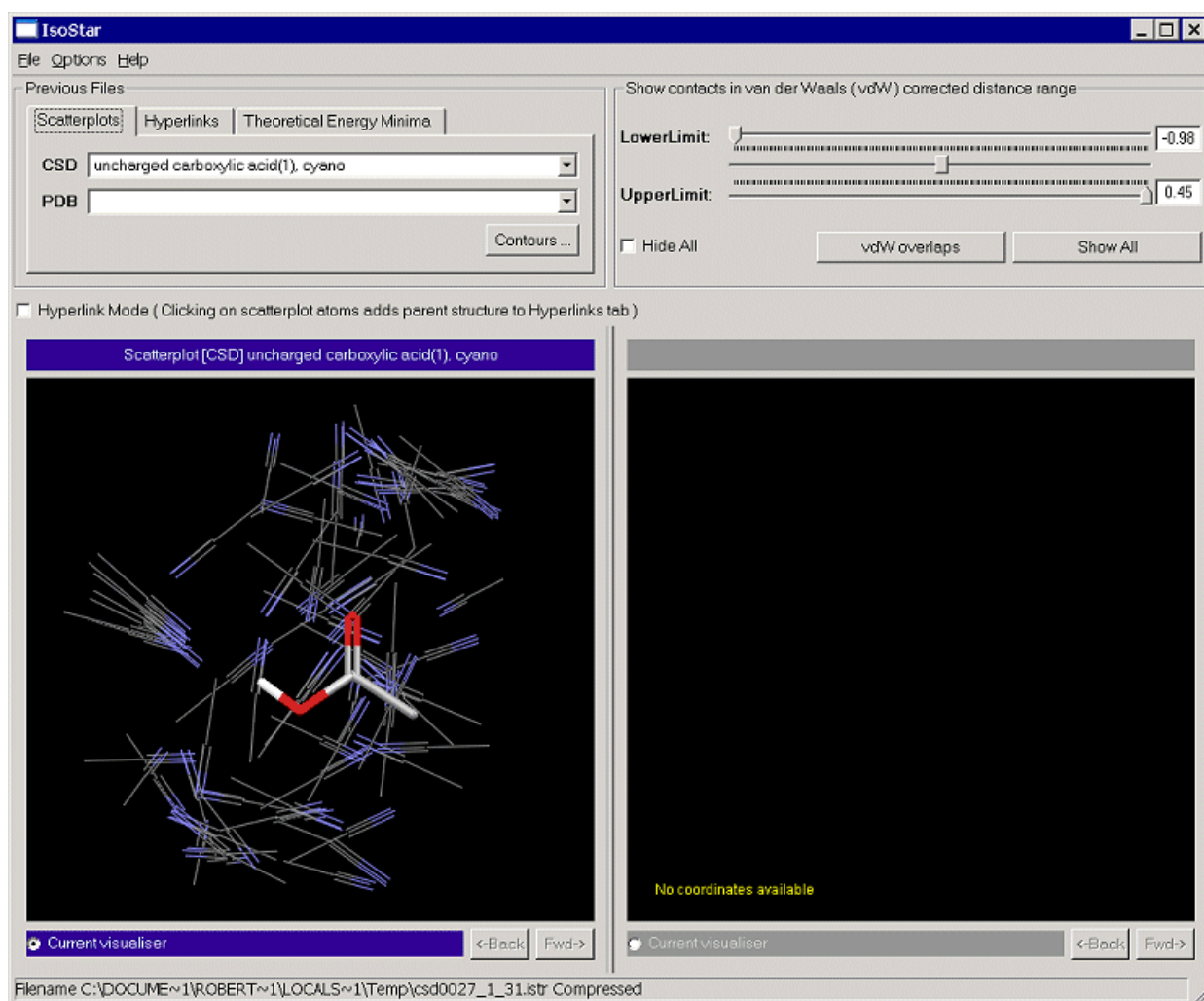
13.4 Selecting the Contact Group

- Once the central group has been selected, you will be presented with a list of available contact groups.

Other N or O				
 Links to statistical data	 Links to theoretical energy data			
Contact Group	CSD	PDB	Stats	Theory
aromatic or sp ² N	1438	367		
cyano	72			
pyramidal R3N	63	*		
uncharged planar R3N	162	56		
cationic planar R3N	16	20		
any terminal O	2488	2300		
any C=O	2492	2300		
amide C=O	551	975		
ester C=O	225	1		
carboxylate	1248	140		
aliphatic ether	158			
linking O in esters	111			
aromatic O	53			
sulfoxide/sulfone O	142			
nitro	516	*		

* This search has not been done.

- The contact groups are listed based on their elemental content.
- The numbers in the CSD or PDB columns correspond to the number of observations of the selected contact in the CSD or PDB respectively.
- Click on the number hyperlink (e.g. [72](#) adjacent to *cyano* under *CSD*) to view the corresponding scatterplot.



- Further information on interacting with IsoStar scatterplots is provided in the IsoStar documentation: <http://www.ccdc.cam.ac.uk/support/documentation/#isostar>.

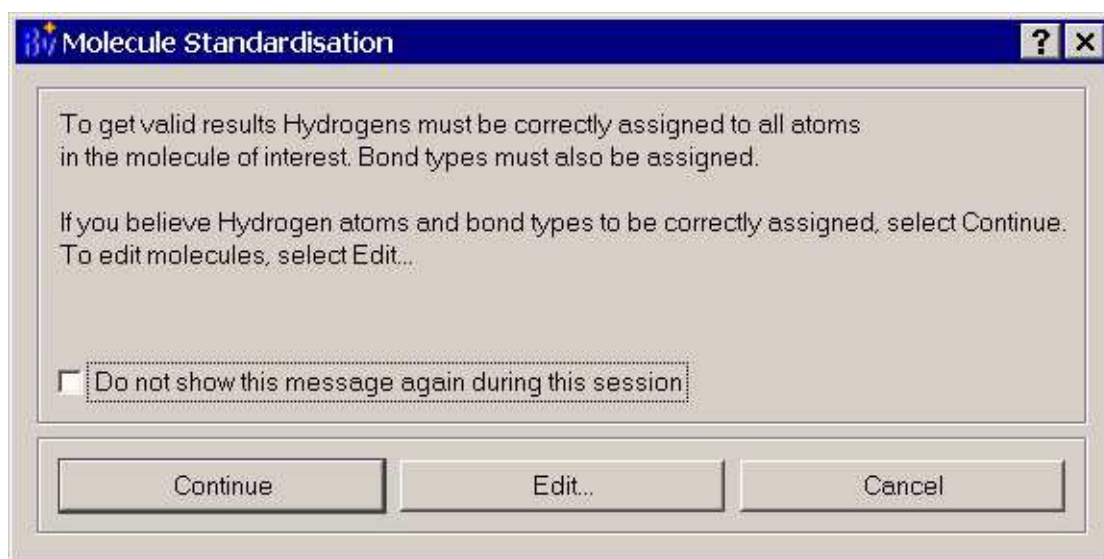
14 Validating Ligand Geometry

14.1 Overview

- Mogul is a library of molecular geometry derived from the Cambridge Structural Database (CSD) and provides information on the preferred values of bond lengths, valence angles and acyclic torsion angles.
- Mogul searches, initiated from within Hermes, can be used e.g. to validate the intramolecular geometries of ligands and identify unusual geometric features.
- Searches can be performed on a specific bond, valence angle or torsion in a query ligand. Alternatively, all geometric features within the loaded ligand(s) can be validated.
- Mogul provides an analysis of the geometry of selected fragments. For example, if you run a search on a bond, valence angle or torsion in a ligand, Mogul will retrieve all entries from the CSD that have the same molecular feature (i.e. bond, angle or torsion fragment within the same structural environment) and return the experimental measurements of those matching fragments, together with the mean and standard deviation of their distribution.
- Query fragments will then be classified according to whether Mogul considers their geometry to be *Unusual* or *Not Unusual* based on statistical measures calculated from matching fragments in the CSD.
- The Mogul geometry checking functionality is only available to registered Cambridge Structural Database (CSD) System subscribers. For further information on obtaining the CSD, please see: http://www.ccdc.cam.ac.uk/contact/obtaining_products/.
- For more information on Mogul, please visit http://www.ccdc.cam.ac.uk/products/csd_system/mogul/.

14.2 Molecule Standardisation

- In order to obtain valid results from Mogul searches it is essential that query ligands are prepared properly. When preparing query structures it is critical that:
 - All ligand hydrogen atoms be included. You should ensure that any hydrogen atoms added reflect the desired ionisation state of the molecule or ion. For example, the protonation state of a carboxylic acid group can be controlled by adding or removing the ionisable hydrogen atom and this will affect the results. The exact 3D positions of any added hydrogen atoms are not important, but the hydrogen atoms must be bonded to the correct heavy atoms.
 - Atom and bond types be defined unambiguously.
 - Correct conventions for bond types are used. Mogul bond types follow the bond type conventions used in the Cambridge Structural Database (CSD). It is critical that these bond type conventions are followed in query structures, otherwise relevant hits may be missed.
- When initiating a Mogul search you will, by default, be presented with the following dialogue:



- If you believe that all ligand hydrogen atoms and bond types are correctly assigned, then select **Continue** to proceed with the Mogul geometry check.
- To have Hermes automatically assign bond types, add missing hydrogen atoms and/or standardise the bond types to CSD conventions, or to manually edit the structure select **Edit** (see Editing Structures, page 93).

Note: There is no guarantee that the automatic bond-type deduction or hydrogen-atom addition will be completely correct, particularly if the query structure has a very poor geometry. Consequently, you should check the automatic changes that are made and if necessary manually edit the structure (see Manually Editing a Structure, page 94).

14.3 Performing a Geometry Check

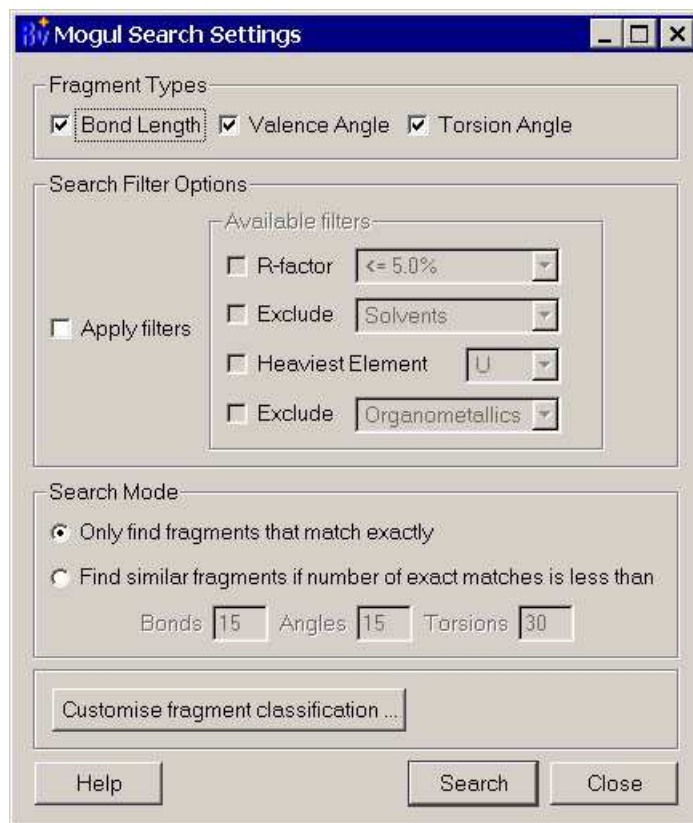
- In order to obtain valid results from Mogul searches it is essential that query ligands are prepared properly (see Molecule Standardisation, page 65).
- Select the atoms that are needed to define the geometric parameter(s) of interest by clicking on them with the left-hand mouse button (see Selecting and Deselecting Atoms and Molecules, page 23). For example, selecting three bonded atoms A-B-C will result in searches being run on the valence angle A-B-C and the two bond lengths A-B and B-C (searches can subsequently be restricted to certain fragment types only).

- If no atoms are selected then a complete analysis of all loaded ligands will be performed.

Note: It is not possible to search for cyclic torsions, i.e. the central bond of the torsion angle must be acyclic. It is not possible to search for bonds, angles or torsions in which one or more defining atoms is a metal or hydrogen atom. It is not possible to search for non-bonded contacts or improper torsions, so chemically-bonded atom sequences must be selected in order to form a valid search fragment.

- Click on **Calculate** in the top-level menu and then select **Mogul Geometry Check...** from the

resulting menu. This will open the *Mogul Search Settings* dialogue:



- The following search and display configuration options are available:
 - *Fragment Types*: It is possible to restrict searches to certain fragment types. Switch on the check-boxes corresponding to the fragment types you wish to include.
 - *Search Filter Options*: Filters can be applied e.g. to exclude CSD structures with low experimental precision (see Applying Search Filters, page 67).
 - *Search Mode*: If the number of hit fragments that are structurally identical to the query fragment is insufficient then it is possible to find similar fragments and have these included in the results (see Finding More Hits, page 68).
 - **Customise fragment classification...** click on this button to customise the classification of fragments in the results display (see Configuring Results Display Options, page 69).
- Hit **Search** to accept the current settings and run the search.

14.3.1 Applying Search Filters

- Search results can be filtered e.g. to exclude structures with low experimental precision.
- Switch on the **Apply filters** check-box in the *Mogul Search Settings* dialogue and select the filters you wish to use. Available filters include:
 - **R-factor**: Switch on the R-factor check box and then select one of the three options from the

drop-down list in order to restrict the results to structures with R-factors less than or equal to 5%, 7.5% or 10%.

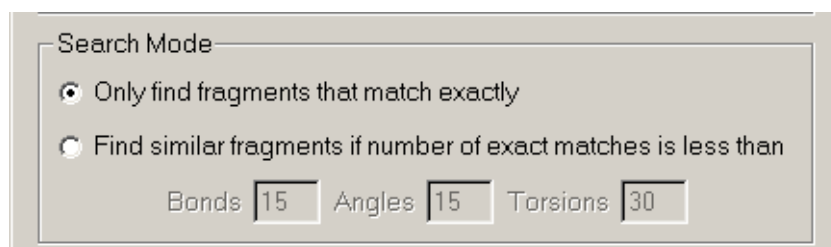
- **Exclusion of Solvents or Non-solvents:** Switch on the Exclude check box and then select either Solvents or Non-solvents from the drop-down list. Results will be filtered to exclude fragments from either solvent or non-solvent molecules.

Note: a molecule is considered to be a solvent if it is found in a CCDC catalogue of known solvent molecules. However, if a crystal structure contains only solvent molecules then the largest molecule will be assumed to be a non-solvent.

- **Heaviest Element:** Switch on the Heaviest Element check box and then select an element from the drop-down list. Results will be filtered to exclude fragments from structures that contain elements heavier than the specified element.
- **Exclusion of Organometallics or Organics:** Switch on the Exclude check box and then select either Organometallics or Organics from the drop-down list. Selecting to exclude Organometallics will eliminate from the results any fragments from structures that contain a transition metal, lanthanide, actinide or any of Al, Ga, In, Tl, Ge, Sn, Pb, Sb, Bi, Po. Selecting to exclude Organics will eliminate fragments from organic structures (i.e. structures that do not contain any of the elements listed above).
- **Note:** when filters have been applied it will not be possible to retrospectively incorporate excluded entries (without running the search again).

14.3.2 Finding More Hits

- When searching on a particular geometric feature, the number of hit fragments found that are structurally identical to the query fragment may not be sufficient. In such cases, the search can look for fragments that, while not identical to the query, are sufficiently similar as to be relevant. These fragments will then be incorporated into the results, which would therefore consists of observations from fragments that are similar to the query fragment as well as fragments that are identical to the query (if these exist).
- To prohibit the inclusion of similar fragments, select the **Only find fragments that match exactly** radio button in the *Mogul Search Settings* dialogue.

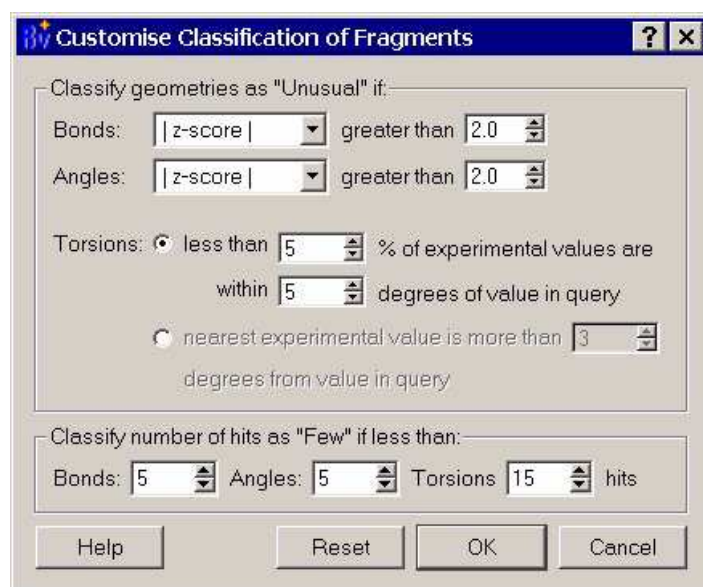


- If a minimum specified number of fragments that are identical to the query are not found a search for similar fragments can be performed. To enable this, select the radio-button labelled

Find similar fragments if number of exact fragments is less than. The criteria for initiating a search for similar fragments can be set independently for bond, angle and torsion angle fragments.

14.3.3 Configuring Results Display Options

- Query fragments will be classified (in the *Mogul Results Viewer*) according to whether Mogul considers their geometry to be *Unusual* or *Not Unusual* based on statistical measures calculated from matching fragments in the CSD.
- To customise the statistical criteria used to classify fragments click on the **Customise fragment classification...** button in the *Mogul Search Settings* dialogue. This will open the *Customise Classification of Fragments* window:

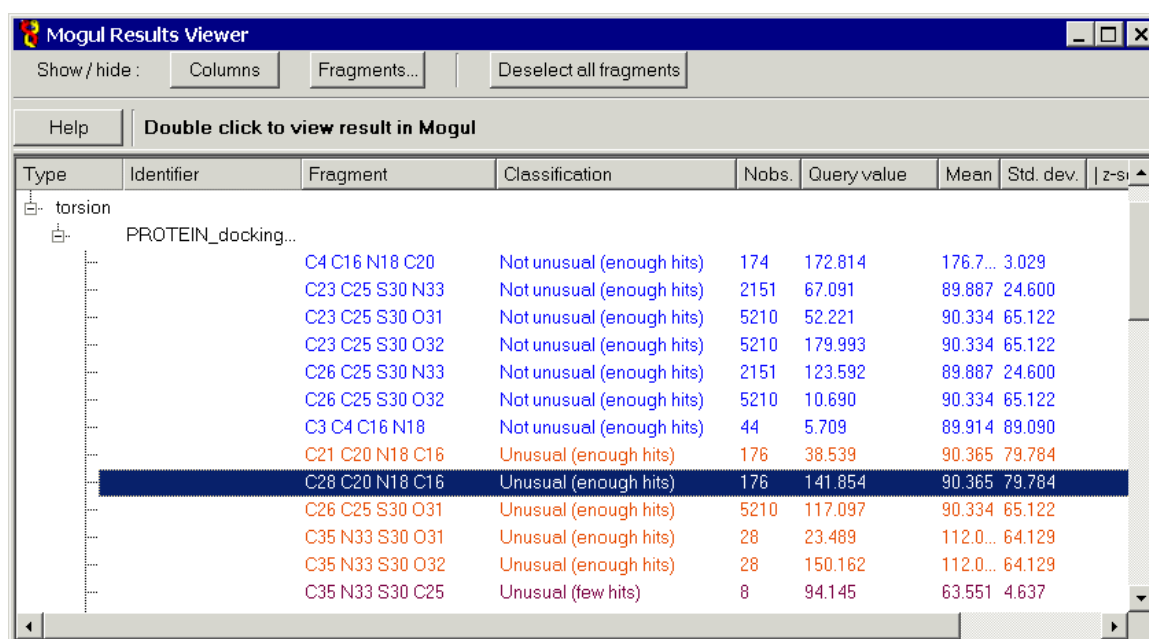


- For bonds and valence angles, the following statistical criteria can be used to classify the geometry of a fragment:
 - $|x - \text{Mean}|$: indicates how far the fragment query value deviates from the mean of the corresponding distribution.
 - $|z\text{-score}|$: is the difference between observed and mean values of a geometric parameter divided by the standard deviation of the distribution.
- For torsion angles, the following statistical criteria can be used:
 - Local Density*: indicates the percentage of experimental measurements that are within x degrees of the value in your query molecule.
 - $|d(\text{min})|$: is the difference between the value of the torsion angle in your query and the nearest torsion angle in the corresponding distribution.
- A bond length, valence angle or torsion angle in your query ligand will subsequently be classified as either *Unusual* or *Not Unusual* according to the criteria specified.

- Query fragments are further classified according to whether or not there is sufficient data to draw a reliable conclusion, i.e. depending on the number of experimental measurements found by Mogul, the data will be classified as *Enough hits* (reliable conclusions can be drawn) or *Few hits* (not enough data to be confident). The criteria for classifying the number of hits as *Few* can be set independently for bond, angle and torsion angle fragments.
- Hit **OK** to accept the current settings, **Cancel** to close the *Customise Classification of Fragments* dialogue without saving any changes, or **Reset** to return settings to their default values.
- The criteria governing the classification of fragments can also be customised after a search has been run. Click on the **Fragments...** button in the *Mogul Results Viewer* followed by **Customise...** in the resulting *Fragment Display Selection* dialogue. By default, changes to classification settings made when viewing results will apply only to the results currently on display. To make the changes apply to all subsequent geometry checks run during the current Mercury session select **Set as default**.

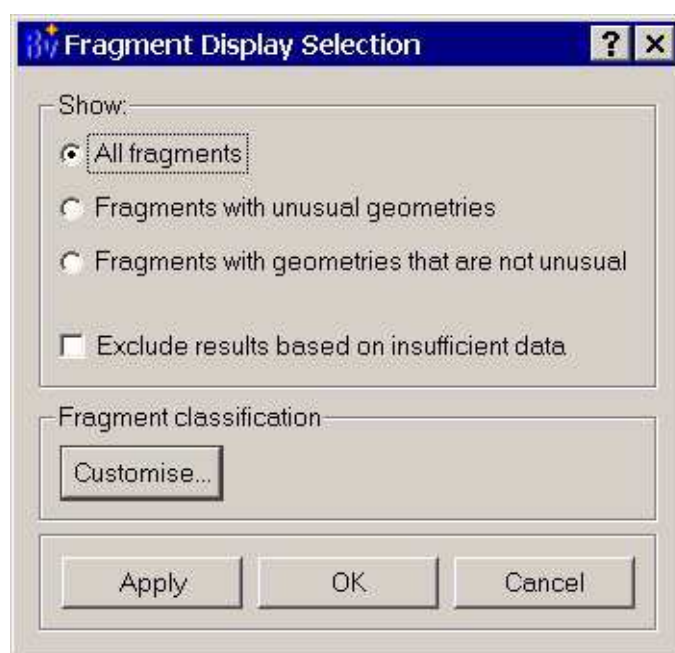
14.4 Viewing and Analysing Results

- On completion of a search, results are displayed in the *Mogul Results Viewer*.
- Data is presented in the form of a tree organised first by fragment type (bond, torsion, angle) and subsequently by structure identifier. Each search fragment is then listed together with its associated statistical measures, calculated from matching entries in the CSD.
- Fragments are classified and colour coded for easy identification of *Unusual* geometric features (see Configuring Results Display Options, page 69). In the example shown below the torsion fragment C28 C20 N18 C16 is classified as *Unusual* when compared to matching entries in the CSD:



Type	Identifier	Fragment	Classification	Nobs.	Query value	Mean	Std. dev.	z-score
torsion	PROTEIN_docking...							
		C4 C16 N18 C20	Not unusual (enough hits)	174	172.814	176.7...	3.029	
		C23 C25 S30 N33	Not unusual (enough hits)	2151	67.091	89.887	24.600	
		C23 C25 S30 O31	Not unusual (enough hits)	5210	52.221	90.334	65.122	
		C23 C25 S30 O32	Not unusual (enough hits)	5210	179.993	90.334	65.122	
		C26 C25 S30 N33	Not unusual (enough hits)	2151	123.592	89.887	24.600	
		C26 C25 S30 O32	Not unusual (enough hits)	5210	10.690	90.334	65.122	
		C3 C4 C16 N18	Not unusual (enough hits)	44	5.709	89.914	89.090	
		C21 C20 N18 C16	Unusual (enough hits)	176	38.539	90.365	79.784	
		C28 C20 N18 C16	Unusual (enough hits)	176	141.854	90.365	79.784	
		C26 C25 S30 O31	Unusual (enough hits)	5210	117.097	90.334	65.122	
		C35 N33 S30 O31	Unusual (enough hits)	28	23.489	112.0...	64.129	
		C35 N33 S30 O32	Unusual (enough hits)	28	150.162	112.0...	64.129	
		C35 N33 S30 C25	Unusual (few hits)	8	94.145	63.551	4.637	

- To control which data items are included in the spreadsheet click on the **Columns** button and select the items you to be displayed: the spreadsheet will update automatically to reflect the selections made.
- The order in which spreadsheet columns are displayed can be changed by pressing down the left-hand mouse button on a column header, moving the cursor while keeping the mouse button depressed to the new location, and then releasing the button.
- The rows of a spreadsheet can be sorted according to the values in any of the columns. For example, to sort the rows by **local-density** (in descending order), click on the **local-density** button at the top of the spreadsheet.
- Clicking on a row in the spreadsheet will result in the chosen fragment being highlighted in the main structure display window. Click on a row again to deselect it. To clear all selections hit the **Deselect all fragments** button.
- Double-clicking on a row will result in the fragment being displayed in the native Mogul interface. From here the underlying histogram and relevant entries from the CSD can be examined in detail.
- It is possible to control which fragments are displayed based on their classification. This can be useful when analysing results from a large number of query fragments. Click on the **Fragments...** button. This will open the *Fragment Display Selection* dialogue:



- Select to display either **All fragments**, **Fragments with unusual geometries**, or **Fragments with geometries that are not unusual**.
- To exclude results for fragments which have *few* hits switch on the **Exclude results based on insufficient data** check-box.
- To customise the statistical criteria used to classify fragments click on the **Customise...** button.

This will open the *Customise Classification of Fragments* window (see Configuring Results Display Options, page 69).

- Hit **Apply** (leaves window open), or **OK** (closes window) to update the *Mogul Results Viewer* according to the selection made. To close the *Fragment Display Selection* window without making any changes hit **Cancel**.

15 Viewing Structures in Stereo

15.1 Overview

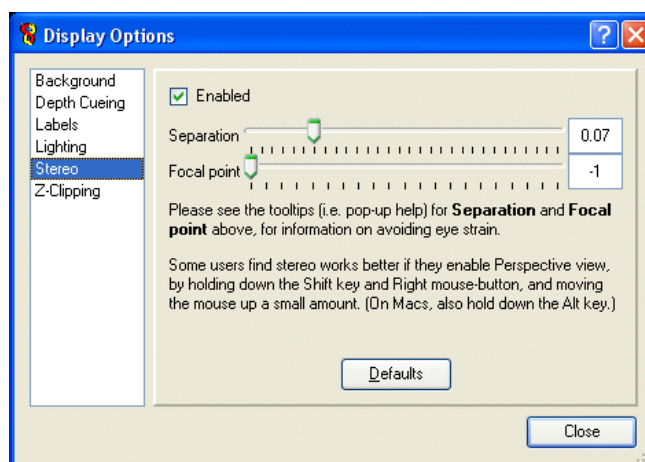
- Stereo viewing works on the premise that most humans use binocular vision to perceive depth and see the world in 3D. Binocular vision relies on the fact that we have two eyes, which are approximately three inches apart. This separation causes each eye to see the world from a slightly different perspective, and the different views are then pieced together by the brain. The brain understands the differences in the two views and uses them to calculate distance, creating our sense of depth and ability to gauge distance.
- To view structures in stereo it is necessary to have the appropriate set-up on your computer, whether this is an emitter with associated stereoviewing eyewear, or a monitor converter with stereoviewing eyewear. Providers of this hardware include Stereographics (<http://www.reald-corporate.com/scientific/products.asp>).
- Only if your computer is correctly configured for stereo viewing will the **Stereo** tickbox, beneath the main menu, be active. If your computer is not configured for stereo viewing or if the settings are not correctly configured, the **Stereo** tickbox will be greyed out.

15.2 Making Use of Stereo Viewing

- If your computer is correctly set-up for stereo viewing, and assuming the hardware required for viewing in stereo has been switched on, stereoscopic viewing can be switched on using the **Stereo** tickbox beneath the main menu options in Hermes.
- Alternatively stereo viewing can be switched on using the **Enabled** tickbox in the *Stereo* window, obtained via **Display**, **Display Settings**.

15.3 Adjusting Stereo Viewing Settings

- The settings for stereo viewing are accessed via **Display** and **Display Settings**, and by clicking on **Stereo** in the resultant window.



- From within this window it is possible to control settings for
 - *Separation*: use the slider to control the amount of separation between the left and right eye views.
Note: Although larger values for *Separation* enhance the stereo view effect, these are more likely to give rise to eye strain and/or headaches.
 - *Focal Point*: use the slider to control how close the focal point is to the front or back of the structure(s) being displayed:
 - 1 - places all structure(s) on display behind the monitor screen
 - 0 - places the middle of the structure(s) in the plane of the screen
 - 1 - places all the structure(s) on display in front of the monitor screen*Note*: Values between -1 and 0 are likely to reduce the effects of eye strain.
- Values for *Separation* and *Focal Point* can also be typed into the boxes to the right of the sliders.
- Default settings can be restored using the **Defaults** button.

15.4 Stereo Viewing: Troubleshooting

- Problems selecting atoms.
 - This can cause problems when viewing in stereo as what you are seeing is the average of two images. To overcome this problem, close one eye (we would recommend the left eye) and select the atom(s) this way.
- The stereo button is disabled in Hermes: stereo can't be turned on.
 - For some reason, the computer thinks that the stereo viewing hardware capability is not available. Check your hardware settings and set-up to ensure the hardware is correctly configured.
- Although stereo settings are turned on and I'm looking through the stereo glasses, I can still see two separate images and not one.
 - Check the emitter is turned on.
 - Check the glasses are turned on.
 - Are the glasses within range of the emitter?
 - Do the batteries in the glasses need replaced?

16 Defining Descriptors

16.1 Overview of Descriptors

- Once a protein has been loaded into the Hermes visualiser, a set of descriptors can be defined. Descriptors might be chosen to quantify as many as possible of the features that you think may correlate with ligand binding affinity, e.g.
 - Ligand properties such as molecular weight and number of rotatable bonds (see Simple Property Descriptors, page 75).
 - Number of protein-ligand hydrogen bonds and bad clashes (see Simple Property Descriptors, page 75).
 - Number of atoms in the protein or ligand that are prevented from fulfilling their hydrogen-bonding capacity (see Simple Property Descriptors, page 75).
 - Presence, absence and type of close contacts to particular atoms in the protein binding site (see General Contact Descriptors, page 78).
 - Presence, absence and geometry of hydrogen bonds to particular atoms in the protein binding site (see Hydrogen Bond Descriptors, page 81).
 - Amount of solvent inaccessible ligand surface area and whether it is composed of hydrophobic or hydrophilic atoms (see Ligand Surface Area Descriptors, page 84).
 - Whether particular regions of the binding site are occupied by ligand atoms (see Sphere Volume Descriptors, page 86).
- The descriptors you define can then be calculated for a library of docked ligands and the results analysed in an external program to identify docking solutions that show promising features (see Specifying the Input to a Descriptor Calculation, page 111).
- The complete set of descriptors (a *descriptor configuration*) can be saved for future use (see Saving Descriptor Definitions, page 89).

16.2 Simple Property Descriptors

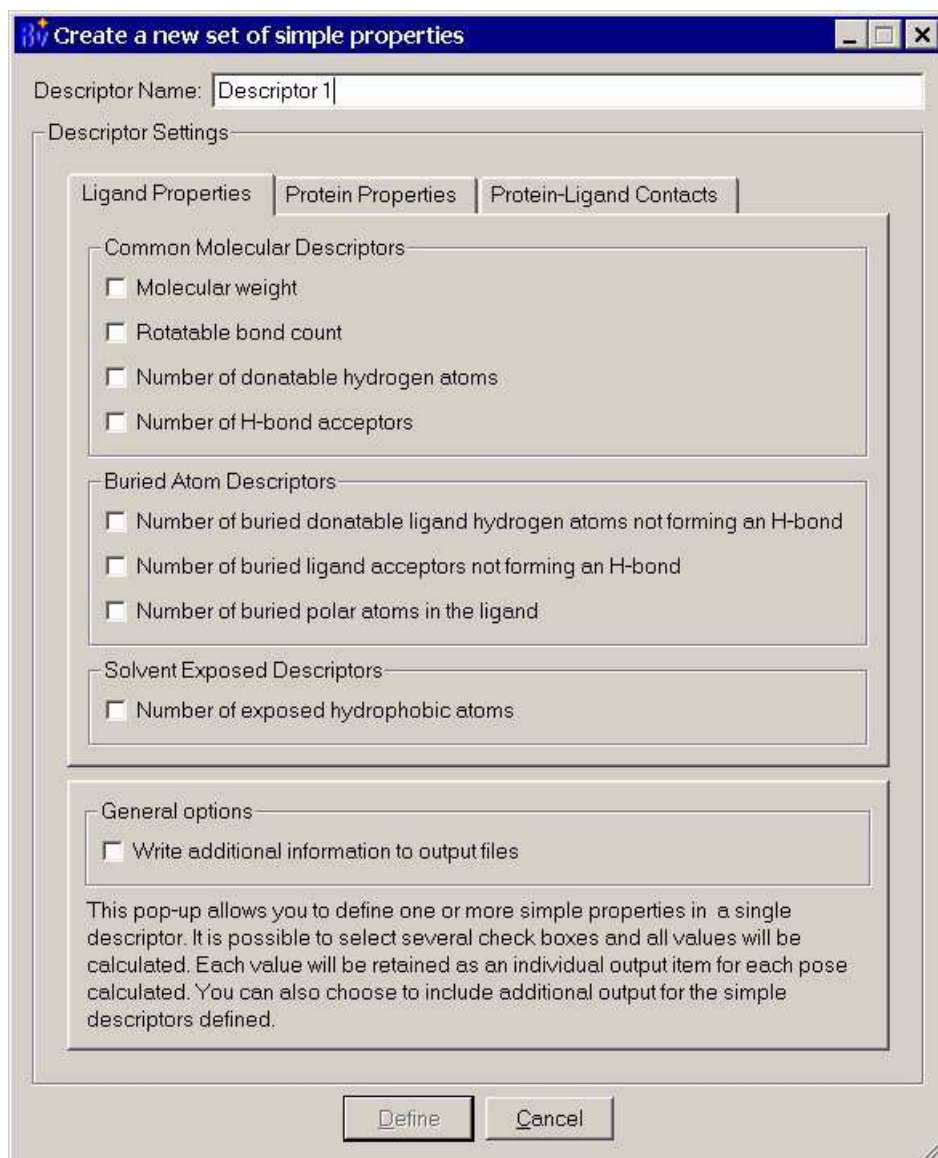
- A variety of simple descriptors that quantify basic properties of the ligand, the protein binding site and the protein-ligand interactions are available. You can select which of these properties you want to be calculated (see Selecting Simple Property Descriptors, page 76).
- Ligand properties include:
 - *Molecular weight*.
 - *Rotatable bond count* (see Rotatable Bonds, page 119).
 - *Number of donatable hydrogen atoms*: The number of ligand hydrogen atoms that are capable of being donated in hydrogen bonds (see Donatable (or Polar) Hydrogens, page 118). Note that it is the hydrogen atoms that are counted, not the donor atoms (e.g. $\text{CH}_3\text{-NH}_2$ would return a value of 2, not 1).

- *Number of H-bond acceptors*: This takes no account of the number of available lone pairs. For example, both $\text{CH}_3\text{-NH}_2$ and $(\text{CH}_3)_2\text{C=O}$ would have an acceptor count of 1, even though the carbonyl oxygen of acetone has two lone pairs available for accepting H-bonds but the methylamine nitrogen has only one.
- *Number of buried donatable ligand hydrogen atoms not forming an H-bond*: The number of ligand polar hydrogen atoms that are prevented from forming hydrogen bonds, e.g. a ligand N-H hydrogen that is desolvated by protein carbon atoms (see Number of Buried Hydrogens/Acceptors Not Forming an H-Bond, page 118).
- *Number of buried ligand acceptors not forming an H-bond*: The number of hydrogen-bond acceptor atoms that are not solvent exposed but also are not hydrogen-bonded to the protein, e.g. a ligand carbonyl oxygen that is in contact with a protein carbon atom (see Number of Buried Hydrogens/Acceptors Not Forming an H-Bond, page 118).
- *Number of buried polar atoms in the ligand*: The number of ligand atoms that are polar but not capable of forming hydrogen bonds (see Polar, non-Hydrogen-Bonding Atoms, page 119) and that are solvent inaccessible (see Buried Atoms, page 118).
- *Number of exposed hydrophobic atoms*: The number of ligand hydrophobic atoms (see Hydrophobic Atoms, page 118) that are accessible to solvent (see Solvent Accessibility, page 119).
- Protein binding-site properties include:
 - *Number of buried donatable protein hydrogens not forming an H-bond*: The number of protein polar hydrogens that are prevented from forming hydrogen bonds by the presence of the ligand (see Number of Buried Hydrogens/Acceptors Not Forming an H-Bond, page 118).
 - *Number of buried protein acceptors not forming an H-bond*: The number of protein hydrogen-bond acceptor atoms that are prevented from forming hydrogen bonds by the presence of the ligand (see Number of Buried Hydrogens/Acceptors Not Forming an H-Bond, page 118).
 - *Number of buried polar atoms in the protein*: The number of atoms that are polar but not capable of forming hydrogen bonds (see Polar, non-Hydrogen-Bonding Atoms, page 119) and that are not solvent exposed (see Buried Atoms, page 118).
- Protein-ligand properties include:
 - *Number of ligand atoms forming H-bonds*: The total number of ligand atoms (polar hydrogens and H-bond acceptor atoms) that are involved in hydrogen bonds to the protein (see Hydrogen Bonds, page 118).
 - *Number of ligand atoms forming bad clashes*: The total number of ligand atoms that clash with protein atoms. A clash is defined as a contact at least $0.5\text{\AA}$ shorter than the sum of van der Waals radii.

16.2.1 Selecting Simple Property Descriptors

- To define a simple property descriptor, click on the top-level menu button **Descriptors**. From the

resulting pull-down menus, select **Define...** followed by **Add Simple Descriptor...**. This will open the *Create a new set of simple properties* dialogue:



- Click on the properties that you want (use the tabs to switch between display of ligand, protein and protein-ligand properties).
- The *Descriptor Name* entry box at the top of the dialogue box will contain a default name; you can type in a name of your own choice if you wish. This will be used as a "root name" for each of the properties you have selected.
- If the *Write additional information to output files* is switched on, the numbers and labels of atoms involved in the various property counts you have selected (e.g. the labels of ligand acceptor atoms if you selected the *Number of H-bond acceptors* descriptor) will be output.
- Hit **Define** to define the descriptor or **Cancel** to abort.

16.3 General Contact Descriptors

- A general contact descriptor can be used to characterise the contacts formed by one or more nominated atoms in the protein to particular types of atoms in the ligand.
- The descriptor can be set up so that it returns the number of contacts formed by each of the chosen protein atoms, or full details of those contacts, or simply whether the protein atoms do or do not form contacts to the ligand (see Controlling the Information Written Out by a General Contact Descriptor, page 80).
- The geometrical criteria defining what constitutes a contact can be customised for each individual general-contact descriptor (see Customising the Geometrical Definition of a Contact, page 80).

16.3.1 Defining a General Contact Descriptor

- To define a general contact descriptor, click on the top-level menu button **Descriptors**. From the resulting pull-down menus, select **Define...** followed by **Add General Contact Descriptor...** This will open the *Create a new general contact descriptor* dialogue:

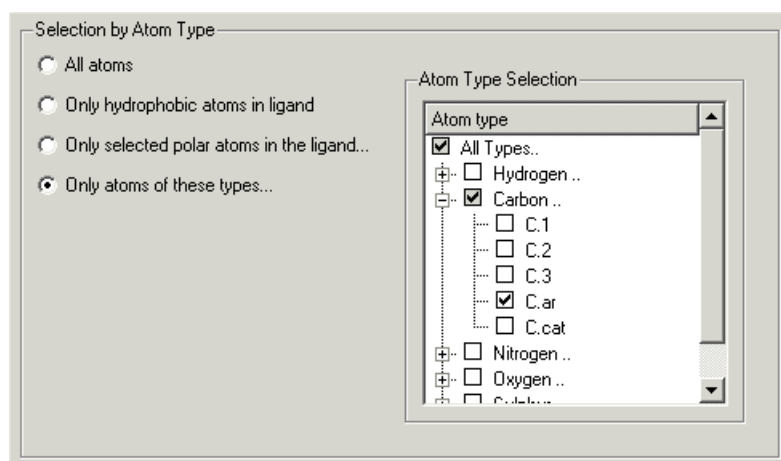
The dialog box is titled "Create a new general contact descriptor". It contains the following fields and options:

- Descriptor Name:** A text field containing "Descriptor 1".
- Descriptor Settings:**
 - Protein subset ...**
 - ☒ All atoms
 - ☐ Atoms in subset [dropdown menu] [Define new..]
 - Ligand types...**
 - Ligand atoms of type [All atoms dropdown menu] [Other...]
 - For each selected protein atom, report**
 - ☒ If it forms a contact to any selected ligand atom
 - ☐ Number of contacts formed to selected ligand atoms
 - ☐ Details of contacts found to selected ligand atoms
 - Output**
 - ☒ Only create output columns for protein atoms that form contacts
 - ☐ Create output columns for all atoms in the protein selection
 - ☐ Combine counts together into a single column
 - Contact geometry options:** [Edit button]
- Buttons:** [Define] [Cancel]

- Nominate the protein atoms and the types of ligand atoms that you want to include in the descriptor (see Specifying the Atoms to be Used in a General Contact Descriptor, page 79).
- When the descriptor is calculated, all contacts between the nominated protein atoms and ligand atoms of the specified type will be found. Select the type of information you want written out for each of the nominated protein atoms (see Controlling the Information Written Out by a General Contact Descriptor, page 80).
- The geometrical definition of a contact (maximum allowed distance, etc.) can be customised if desired (see Customising the Geometrical Definition of a Contact, page 80). The customised definition will apply only to the descriptor for which it was created.
- A default name will be assigned to the descriptor; use the *Descriptor Name* entry box if you want to specify your own.
- Hit **Define** to define the descriptor or **Cancel** to abort.

16.3.2 Specifying the Atoms to be Used in a General Contact Descriptor

- If you want information on contacts formed (to the ligand) by *all* protein atoms, switch on the *All atoms* option.
Warning: this might generate a lot of output.
- If you want information on contacts formed by particular protein atoms, you must do this by creating a protein subset (see Making Complex Selections: Defining Subsets of Protein Atoms, page 24) containing just those atoms. Switch on the *Atoms in subset* option, hit **Define new...** and construct the required protein subset (see Creating Protein Subsets, page 24). If the required protein subset already exists, pick it from the selection box that is located to the left of the Define new... button.
- Use the selection box labelled *Ligand atoms of type* to specify which types of ligand atoms are to be considered, e.g. *All hydrophobic atoms*. Hitting the **Other...** button opens a dialogue box that allows a higher degree of control over which ligand atoms are included. In particular, selecting the *Only atoms of these types...* option in this dialogue box allows you to select atoms according to their element or atom type (see Atom Types, page 117). Subdivision by element type is shown by default; clicking on a "+" symbol in the *Atom Type Selection* area displays a further subdivision by atom type. Element or atom types opposite ticked check boxes will be included in the calculation:



16.3.3 Controlling the Information Written Out by a General Contact Descriptor

- When a general contact descriptor is calculated, information will be written out for each of the protein atoms that has been included in the descriptor definition (see Specifying the Atom(s) to be Used in a Hydrogen Bond Descriptor, page 82).
- The nature of this information is controlled by the three radio buttons underneath the words *For each selected protein atom, report:*
 - Switch on *If it forms a contact to any selected ligand atom* if you just want to know whether each protein atom is or is not forming at least one contact to ligand atoms of the specified type(s).
 - Switch on *Number of contacts formed to selected ligand atoms* if you want to know how many contacts each protein atom forms.
 - Switch on *Details of contacts found to selected ligand atoms* if you just want full details of the hydrogen bonds formed by each protein atom. The contact distance, ligand atom index and ligand atom label will be given in that order.

16.3.4 Customising the Geometrical Definition of a Contact

- Click on the *Contact geometry options* **Edit** button.
- In the resulting dialogue, use the radio buttons to specify whether the minimum and maximum allowed contact distances are to be specified in absolute terms (Å) or relative to the van der Waals radii of the interacting atoms.
- Use the entry boxes to specify the minimum and maximum allowed distances. An interaction will not be considered a contact if it falls outside the specified limits.
- For example, selecting the *Actual distance* option and specifying minimum and maximum distances of 1.0 and 3.0 respectively will find contacts between 1 and 3 Å. Selecting *Vdw corrected distance* and typing in minimum and maximum values of -2.0 and -0.5 will find all contacts in the range (S - 2.0) and (S - 0.5), where S is the sum of the van der Waals radii of the

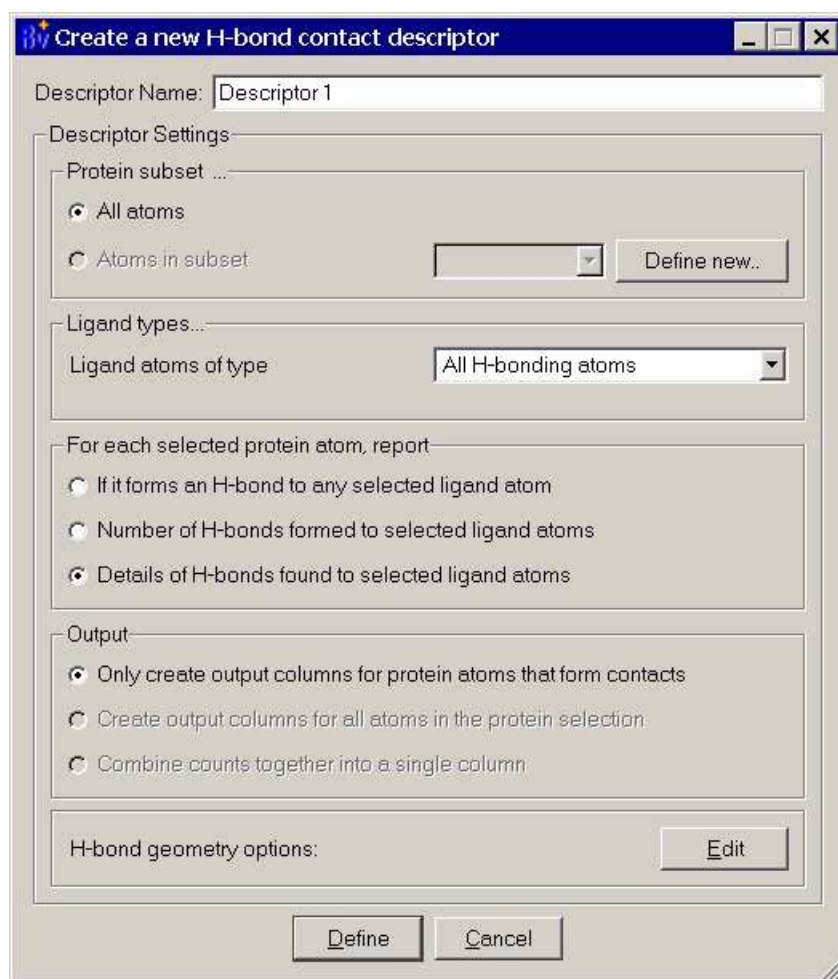
interacting atoms.

16.4 Hydrogen Bond Descriptors

- A hydrogen-bond descriptor can be used to characterise the H-bonding interactions formed by one or more nominated atoms in the protein to particular types of atoms in the ligand.
- A hydrogen bond descriptor can be set up so that it returns the number of H-bonds formed by each of the nominated protein atoms, or full details of those H-bonds, or simply whether the protein atoms are or are not H-bonded to the ligand (see Controlling the Information Written Out by a Hydrogen Bond Descriptor, page 83).
- The geometrical criteria defining what constitutes a hydrogen bond can be customised for each individual hydrogen-bond descriptor (see Customising the Geometrical Definition of a Hydrogen Bond, page 83).

16.4.1 Defining a Hydrogen Bond Descriptor

- To define a hydrogen bond descriptor, click on the top-level menu button **Descriptors**. From the resulting pull-down menus, select **Define...** followed by **Add H-bond Descriptor**. This will open the *Create a new H-bond contact descriptor* dialogue:



- Nominate the protein atoms and the types of ligand atoms that you want to include in the descriptor (see Specifying the Atom(s) to be Used in a Hydrogen Bond Descriptor, page 82).
- When the descriptor is calculated, all hydrogen bonds between the nominated protein atoms and ligand atoms of the specified type will be found. Select the type of information you want written out for each of the nominated protein atoms (see Controlling the Information Written Out by a Hydrogen Bond Descriptor, page 83).
- The geometrical definition of a hydrogen bond (maximum allowed distance, etc.) can be customised if desired (see Customising the Geometrical Definition of a Hydrogen Bond, page 83). The customised definition will apply only to the descriptor for which it was created.
- A default name will be assigned to the descriptor; use the *Descriptor Name* entry box if you want to specify your own.
- Hit **Define** to define the descriptor or **Cancel** to abort.

16.4.2 Specifying the Atom(s) to be Used in a Hydrogen Bond Descriptor

- If you want information on hydrogen bonds formed to the ligand by *all* protein atoms, switch on the *All atoms* option.

Warning: this might generate a lot of output.

- If you want information on hydrogen bonds formed by particular protein atoms, you must do this by creating a protein subset (see Making Complex Selections: Defining Subsets of Protein Atoms, page 24) containing just those atoms. Switch on the *Atoms in subset* option, hit **Define new...** and construct the required protein subset (see Creating Protein Subsets, page 24). If the required protein subset already exists, pick it from the selection box that is located to the left of the Define new... button.

Note: if you want to find H-bonds donated by the protein to the ligand, you need to include the polar protein *hydrogen* atoms in the subset, not the donor atoms (e.g. the H atom of an N-H group, not the N atom). You must therefore ensure that hydrogen atoms are included on all H-bond donor atoms in your subset. If a defined subset contains H-bond donor atoms but does not include all the hydrogen atoms to which they are bonded, then the descriptor will not find any hydrogen bonds donated by these donors. When defining a protein subset you will be given the option to add the required hydrogen atoms automatically (see Creating Protein Subsets, page 24).

- Use the *Ligand atoms of type* selection box to specify the types of ligand atoms that you wish to be considered, i.e. all H-bonding atoms, just donor hydrogens, or just acceptors. Depending on the protein atoms you have chosen, some of these may be not be sensible choices. For example, if you have specified a protein subset containing no acceptor atoms, there is no point in specifying ligand atoms of type *All H-bond hydrogens*.

16.4.3 Controlling the Information Written Out by a Hydrogen Bond Descriptor

- When a hydrogen bond descriptor is calculated, information will be written out for each of the protein atoms that has been included in the descriptor definition (see Specifying the Atom(s) to be Used in a Hydrogen Bond Descriptor, page 82).
- The nature of this information is controlled by the three radio buttons underneath the words *For each selected protein atom, report:*
 - Switch on *If it forms an H-bond to any selected ligand atom* if you just want to know whether each protein atom is or is not forming at least one hydrogen bond to ligand atoms of the specified type(s).
 - Switch on *Number of H-bonds formed to selected ligand atoms* if you want to know how many H-bonds each protein atom forms.
 - Switch on *Details of H-bonds formed to selected ligand atoms* if you just want full details of the hydrogen bonds formed by each protein atom. The H-bond (H...A) distance, ligand atom index and ligand atom label will be given in that order.

16.4.4 Customising the Geometrical Definition of a Hydrogen Bond

- Click on the *H-bond geometry options Edit* button.
- In the resulting dialogue, the *Require hydrogen atom to be present* check box cannot be switched

off. This means that Hermes will assume that protein and ligand hydrogen atoms are present and in meaningful positions and will use the H-atom positions in deciding whether a given interaction is or is not an H-bond.

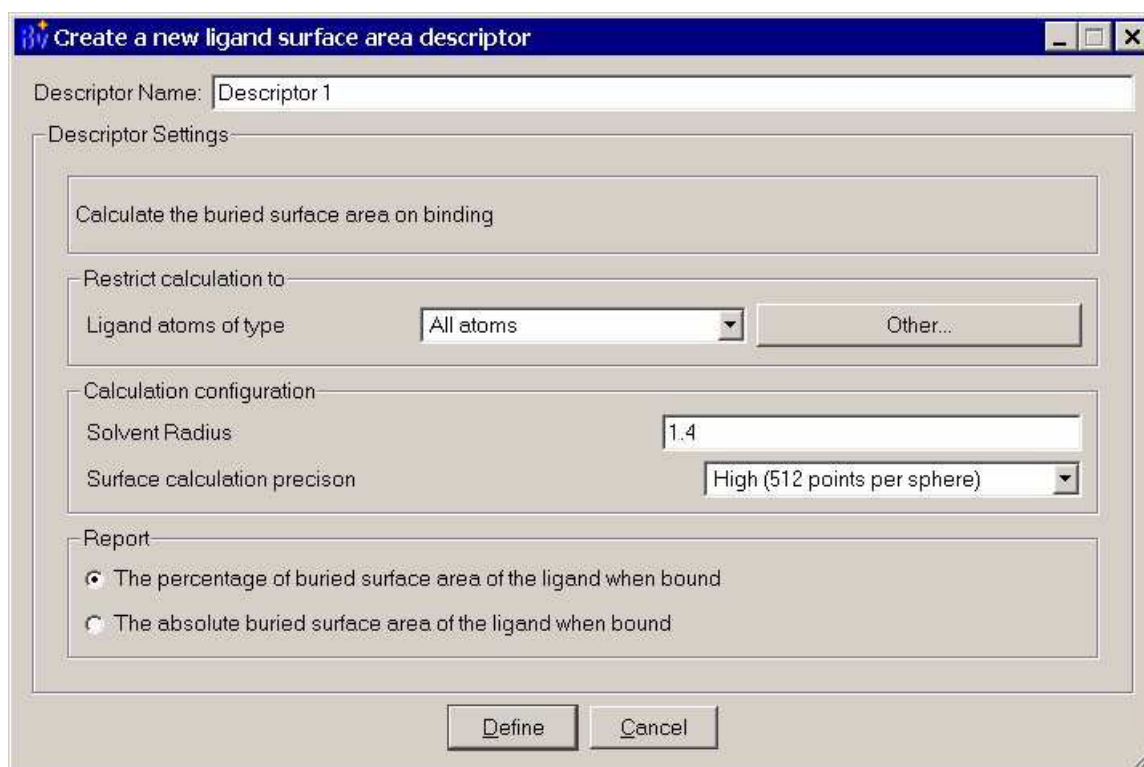
- Use the *D-H...A angle* $\geq$ entry box to specify the minimum allowed donor-hydrogen...acceptor angle. An interaction will not be considered a hydrogen bond if the angle falls below this value.
- Use the radio buttons at the bottom of the dialogue box to specify whether the minimum and maximum allowed hydrogen-bond distances are to be specified in absolute terms (Å) or relative to the van der Waals radii of the interacting atoms.
- Use the entry boxes to specify the minimum and maximum allowed H...acceptor distances. An interaction will not be considered a hydrogen bond if it falls outside the specified limits.
- For example, selecting the *Actual distance* option and specifying minimum and maximum distances of 1.0 and 2.0 respectively will find H-bonds whose H...acceptor distances fall between 1 and 2Å. Selecting *Vdw corrected distance* and typing in minimum and maximum values of -2.0 and -0.5 will find all H-bonds with H...A distances in the range (S - 2.0) and (S - 0.5), where S is the sum of the van der Waals radii of the H and acceptor atoms.

16.5 Ligand Surface Area Descriptors

- A ligand surface area descriptor measures how much of a ligand is solvent inaccessible before and after docking (see Solvent Inaccessible Ligand Surface Area, page 119). Obviously, high-affinity ligands may normally be expected to have a large contact surface area with the protein, so this may often be one of the more discriminating descriptors.
- The calculation can be restricted to particular types of ligand atoms. A common choice, for example, may be to confine the calculation to hydrophobic ligand atoms, since burial of hydrophobic ligand groups is often very important to activity.
- The solvent-inaccessible surface area can be expressed in absolute terms (Å²) or as a percentage of the total surface area of the ligand (or, e.g., of hydrophobic ligand atoms if the calculation was restricted to atoms of this type).

16.5.1 Defining a Ligand Surface Area Descriptor

- To define a ligand surface area descriptor, click on the top-level menu button **Descriptors**. From the resulting pull-down menus, select **Define...** followed by **Add Ligand Surface Area Descriptor...** This will open the *Create a new ligand surface area descriptor* dialogue:



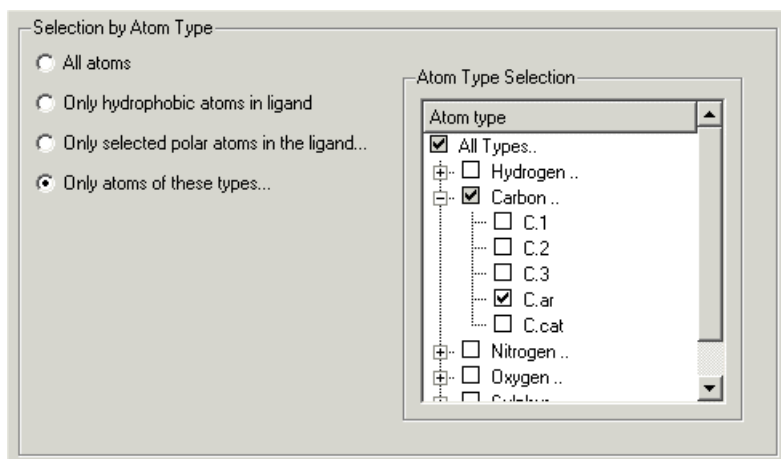
- Specify the types of ligand atoms that you want to include in the descriptor (see Specifying the Atoms to be Included in Calculating a Ligand Surface Area Descriptor, page 85).
- Use the *Calculation configuration* panel to define the *Solvent Radius* and/or change the *Surface calculation precision* settings. These settings allow the speed of the surface area descriptor calculation to be configured, i.e. you can reduce the solvent radius or the number of surface points to speed up the calculation.
- Use the radio buttons to specify whether the buried surface area should be reported in absolute terms (A^2) or as a percentage of the total surface area of the ligand (see Solvent Inaccessible Ligand Surface Area, page 119).
- A default name will be assigned to the descriptor; use the *Descriptor Name* entry box if you want to specify your own.
- Hit **Define** to define the descriptor or **Cancel** to abort.

16.5.2 Specifying the Atoms to be Included in Calculating a Ligand Surface Area Descriptor

- By default, all ligand atoms will be included, so the descriptor will return the total surface area of the ligand that is solvent inaccessible (or the percentage of that total surface area).
 - Two numbers will actually be returned, corresponding to the ligand surface area before and after docking (see Solvent Inaccessible Ligand Surface Area, page 119).
- You may well want to restrict the calculation to particular types of ligand atoms; this can be done by changing the *Ligand atoms of type* selection box to, e.g., *All hydrophobic atoms*. The

descriptor will now return how much of the hydrophobic ligand surface area is solvent inaccessible.

- Hitting the **Other...** button opens a dialogue box that allows a higher degree of control over which ligand atoms are included. Of particular note in this dialogue box are:
 - Selecting the *Only atoms of these types...* option allows you to select atoms according to their element or atom type (see Atom Types, page 117). Subdivision by element type is shown by default; clicking on a "+" symbol in the *Atom Type Selection* area displays a further subdivision by atom type. Element or atom types opposite ticked check boxes will be included in the calculation:



- Selecting the *Only selected polar atoms in the ligand...* allows you to restrict the calculation to a particular type of polar atom, e.g. H-bonding acceptor atoms.

16.6 Sphere Volume Descriptors

- A sphere volume descriptor can be used to determine whether a particular region of the binding site is occupied by ligand atoms. For example, you might want to do this if you believe that molecules are unlikely to be active unless they fill a specific pocket.
- One or more spheres are placed in the region of interest. Each is assigned a radius, so some spheres can be bigger than others, if required. Thus, the spheres can be adjusted, e.g, to fill an entire pocket in the binding-site.
- The descriptor measures how much of the volume of the sphere(s) is accessible (i.e. unoccupied by atoms) before and after docking of the ligand (see Sphere Accessible Volume, page 119).
- The unoccupied surface volume can be expressed in absolute terms (A^2) or as a percentage of the total volume of the sphere(s).
- The descriptor is relatively time-consuming to calculate, but you can control the balance between the speed and accuracy of volume estimation.

16.6.1 Defining a Sphere Volume Descriptor

- To define a sphere volume descriptor, click on the top-level menu button **Descriptors**. From the resulting pull-down menus, select **Define...** followed by **Add Sphere Volume Descriptor...** This will open the *Create a new sphere accessible volume descriptor* dialogue:

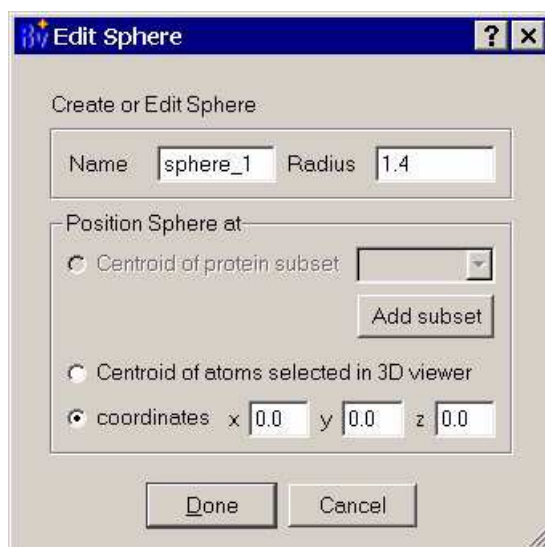
The dialog box is titled "Create a new sphere accessible volume descriptor". It contains several sections:

- Descriptor Name:** A text box containing "Descriptor 1".
- Descriptor Settings:**
 - A checkbox labeled "Save accessible volume before and after docking of a set of spheres" is checked.
- Defined points:**
 - A table with columns: Name, x, y, z, Radius.
 - Below the table are three buttons: "Add New Sphere", "Edit Selected", and "Delete Selected".
- Calculation Options:**
 - A label "Sampling density (number of samples per cubic Å)" followed by a spin box set to "25".
 - A label "Ligand atoms of type" followed by a dropdown menu set to "All atoms" and an "Other..." button.
- Save Options:**
 - Two radio buttons:
 - ☐ Save percentage of accessible sphere volume when the ligand binds
 - ☒ Save absolute amount of accessible sphere volume when the ligand binds
- Buttons:** "Define" and "Cancel" at the bottom.

- Create the spheres that you want to use (see Creating Spheres, page 88).
- When you have created all the spheres you want, use the radio buttons at the bottom of the *Create a new sphere accessible volume descriptor* dialogue to specify whether the descriptor should be the absolute or the percentage of volume that is not occupied (see Sphere Accessible Volume, page 119).
- All the spheres in the *Defined spheres* list box will be included in the calculation.
- A default name will be assigned to the descriptor; use the *Descriptor Name* entry box if you want to specify your own.
- Hit **Define** to define the descriptor or **Cancel** to abort.

16.6.2 Creating Spheres

- Click on **Add New Sphere**. This will open the *Edit sphere* dialogue:



- The position of the new sphere can be defined in any of three ways:
 - At the centroid of a protein subset: select the *Centroid of protein subset* option and either create a new protein subset (hit **Add subset**) or, if you have already created the subset you want, pick it using the selection box immediately above the **Add subset** button.
 - At the centroid of manually-selected atoms: select the *Centroid of atoms selected in 3D viewer* option, and click on the required atoms in the 3D visualiser. You may pick atoms in both the protein and any ligand that you may have displayed.
 - At an explicitly-defined position: select the *Coordinates* option and type in the required x,y,z coordinates in the entry boxes.
- Define a name for the new sphere, if desired (a default name will be supplied).
- Change the radius of the sphere, if desired.
- The sphere will be displayed in the visualiser and its position and radius altered in response to any changes you make in the dialogue box.
- When the sphere is as you want it, hit **Done**.
- You can then repeat the procedure to define more spheres. As spheres are created, they will be listed in the white area of the *Create a new sphere accessible volume descriptor* dialogue box.
- To edit a sphere, select it in the white list box and hit **Edit Selected** to re-open the Edit spheres dialogue.
- To delete a sphere, select it in the white list box and hit **Delete Selected**.

16.6.3 Balancing the Speed and Accuracy of Sphere Volume Estimation

- Sphere volume calculation is relatively time consuming. To speed it up (but at the cost of some accuracy), set the *Sampling density* to a lower number within the *Create a new sphere accessible volume descriptor* dialogue.

16.7 Displaying, Editing and Deleting Descriptor Definitions

- To view the descriptors you have created click on the top-level menu button **Descriptors**. Then from the resulting pull-down menus, enable the **Descriptor Browser**. A list of created descriptors will appear in the *Descriptors* pane in the *Protein Explorer* window.
- To examine the details of and/or edit a descriptor, find it in the *Descriptors* list and double-click on it (or click on it with the right-hand mouse button and select **Edit this descriptor...** from the resulting pull-down menu). This will open the relevant descriptor-definition dialogue box.
- To delete a descriptor, find it in the *Descriptors* list, click on it with the right-hand mouse button and select **Remove this descriptor...** from the resulting pull-down menu.

16.8 Saving Descriptor Definitions

- To save the currently-defined descriptors, click on **Descriptors** from the top-level menu, select **Export Descriptor Configuration File...** in the resulting pull-down menu, and then specify a file name in the dialogue box.
- Descriptor configurations are saved in *xml* format. If you do not specify the *xml* file extension, it will be added automatically.

17 Calculating and Using Descriptors

Much of this functionality is specific to other CCDC software packages (see Calculating Descriptors for a Set of Docked Ligands, page 111).

17.1 Using Calculated Descriptors in External Programs

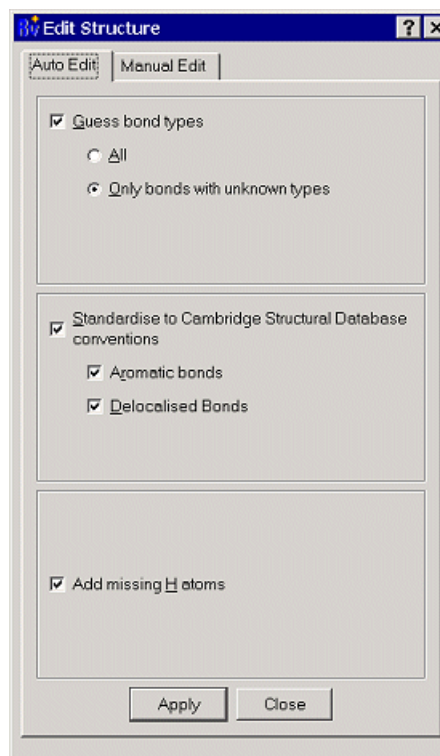
- The results of a descriptor calculation can be written to a tagged *sdf* file or a comma-separated-value (*csv*) file. Each of these formats may be read into various molecular and general-purpose spreadsheet packages.
- For example, a *csv* format file may be read directly into Microsoft Excel.
- Each calculated descriptor value will be named or tagged. The name or tag will include the descriptor name that you specified when creating the descriptor but may have additional characters too.

18 Editing Structures

- A number of structure editing options are available via the **Edit** menu:
 - **Combined Edit Structure**: both the **Edit Structure** and **Auto-Edit Ligand** functionality is accessible from one dialogue via this option (see Combined Manual and Auto-Editing of a Structure, page 93).
 - **Edit Structure** (see Manually Editing a Structure, page 94).
 - **Auto-Edit Ligand** (see Auto-Editing Options, page 95).
 - **Correct Ligand Connectivity** (see Automatically Correcting Ligand Geometry, page 95).

18.1 Combined Manual and Auto-Editing of a Structure

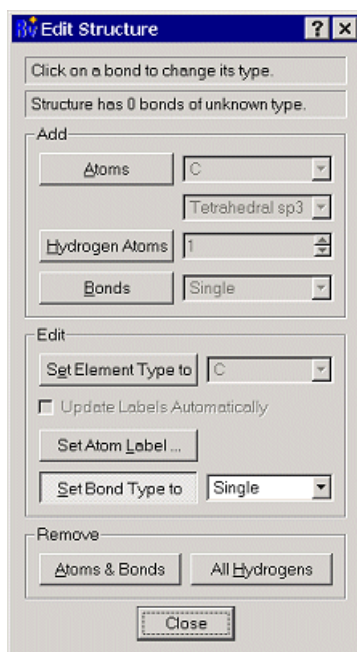
- Combined functionality for manual or auto-editing of structures is accessed by selecting **Edit** then **Combined Edit Structure**.



- The display is controlled using the *Auto Edit* and *Manual Edit* tabs at the top of the window.
- The functionality provided under each tab is described in the sections that follow:
 - *Auto Edit*: See “Auto-Editing Options” on page 95.
 - *Manual Edit*: See “Manually Editing a Structure” on page 94.

18.2 Manually Editing a Structure

- To manually edit a protein or ligand structure, select **Edit** then **Edit Structure**. This will launch the Edit Structure window:



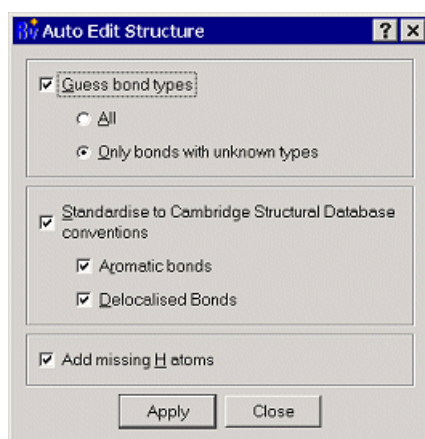
Within this window you can choose to:

- Add atoms: click on the **Atoms** button and select the required atom from the adjacent pull-down menu. If the required atom is not present in the **Atoms** pull-down menu, select the **Other...** option and pick from the periodic table. The hybridisation state of the atom can be selected from the second menu which reads *Tetrahedral sp3* above. Once you have selected the new atom and its hybridisation state, click on the atom(s) to which the new atom should be bonded.
- Add hydrogen atoms: click on the **Hydrogen Atoms** button, select the number of required hydrogen atoms, then click on the atom(s) to which hydrogens are to be added.
- Add bonds: click on the **Bonds** button, choose the bond type from the pull-down menu, then click on the two atoms that you wish to be bonded.
- Modify element type: click on the **Set Element Type to** button, select the new atom type from the pull-down menu adjacent, then click on the atom you would like to modify. Activate the *Update Labels Automatically* tickbox if you require the atom label to be changed with the atom modification.
- Modify an atom label: click on the **Set Atom Label** button and select the atom that requires a new label. Enter the text for the new atom label into the resultant *Edit Atom* window, then click **OK** when the text is correct.
- Edit a bond type: click on the **Set Bond Type to** button, choose the bond type from the pull-down menu, then pick the bond(s) you wish to change.

- Remove atoms and bonds: click on the **Atoms & Bonds** button, then pick on the atoms or bonds you wish to remove.
- Remove all hydrogen atoms: click on the **All Hydrogens** button and select the molecule that H atoms should be removed from.

18.3 Auto-Editing Options

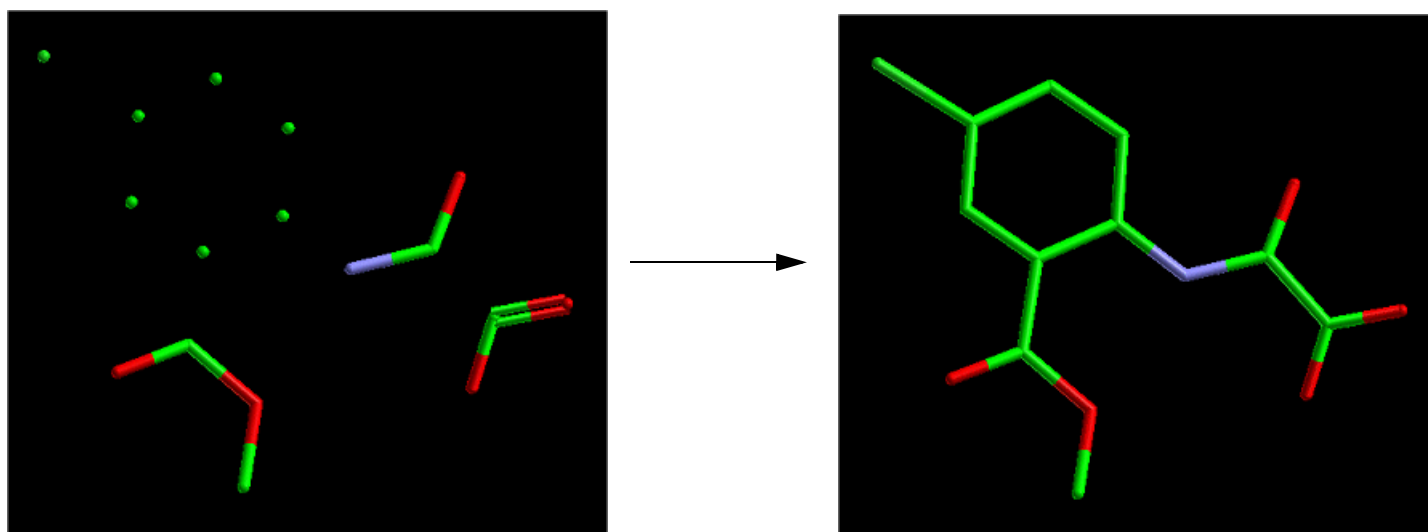
- There are a number of options for editing a ligand structure automatically. These can be accessed by selecting **Auto Edit Ligand** from the **Edit** menu. This will launch an *Auto Edit Structure* window:



- The options available within this window are:
 - Guess bond types: select the **Guess bond types** check box, then apply this to either **All** or **Only bonds with unknown types**.
 - Standardise bond types: select the **Standardise to Cambridge Structural Database conventions** check box. This can be applied to **Aromatic bonds**, **Delocalised bonds** or **Both**.
 - Add missing hydrogen atoms: select the **Add missing H atoms** check box. This is important if using the geometry validating feature available within Hermes (see Validating Ligand Geometry, page 65).
- Hit the **Apply** button to automatically edit your structure according to the selections made.

18.4 Automatically Correcting Ligand Geometry

- In cases where a ligand is badly disordered, the ligand geometry can be corrected by Hermes using the **Correct Ligand Connectivity** option from the **Edit** menu.



- Please note that there is no undo feature for auto-correction; if the results are not satisfactory, you will need to re-load the structure or structure file again and manually edit the structure (see Manually Editing a Structure, page 94).

19 Saving Files

The Hermes visualiser can write:

- Graphics files containing the image currently on display (see Saving Hermes Visualiser Displays as Images, page 97).
- Files containing the 3D coordinates of the molecule(s) on display. You have the choice of saving as:
 - A *.reliarch* file, which retains full information about display properties such as atom colours, etc., but can only be read back into Hermes (see Saving Structures in the Hermes Visualiser Format, page 97).
 - A file in an external formats (*.mol2 .pdb*), which does not hold information about all the Hermes visualiser display settings but can be read into other programs (see Saving Structures in External (3rd Party) Formats, page 98).

19.1 Saving Hermes Visualiser Displays as Images

- The current view can be saved as a graphical image in the following file formats:
 - Bitmap file (**.bmp*)
 - Portable Network Graphics file (**.png*)
 - jpeg files (**.jpg* or **.jpeg*); not available on all platforms.
 - PovRAY files (The Persistence of Vision Raytracer, **.pov*), see <http://www.povray.org/>.
- To save the current image, hit **File** in the top-level menu and **Save As...** in the next menu. In the resulting *Save File As* dialogue box, specify the name of the file to which you want to write, choose the required format in the *Save As* entry box, and hit **Save**.
- The *Save Image* dialogue box then appears. This allows you to specify the image size. Sizes over 1500 by 1500 pixels require large amounts of memory. This dialogue box also allows you to thicken the lines in the image.

19.2 Saving Structures in the Hermes Visualiser Format

- The Hermes visualiser has its own file format (filename extension *.reliarch*), which will store complete information about the structure(s) currently on display, including not only the 3D atomic coordinates but also display properties, etc.
- To save the currently loaded structures and display settings, hit **File** in the top-level menu and **Save As...** in the next menu. In the resulting *Save File As* dialogue box, specify the name of the file to which you want to write, choose **Binary Archive (*.reliarch)** in the *Save as type* entry box, and hit **Save**.
- A *.reliarch* file can only be read back into Hermes, not into other programs.
- The *.reliarch* file is binary so cannot be read using a text editor.

19.3 Saving Structures in External (3rd Party) Formats

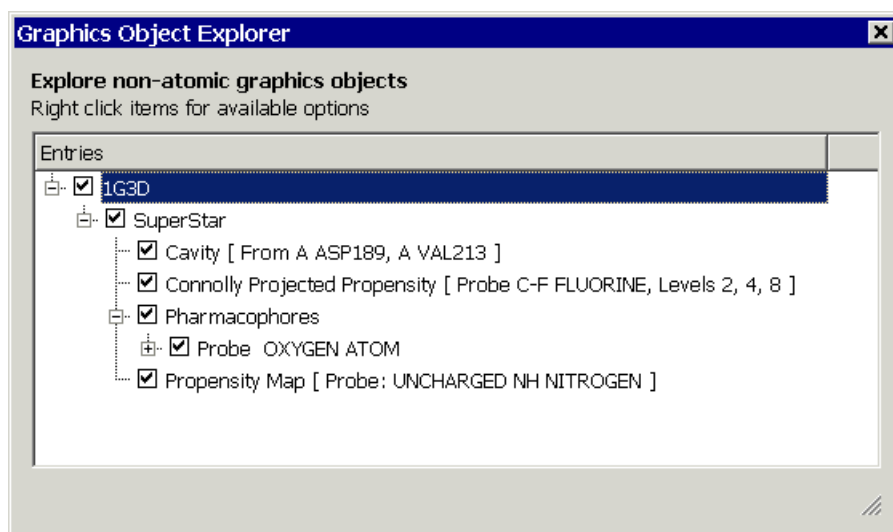
- Hermes can write structures in the following external file formats:
 - *mol2*
 - *pdb*
- To save the currently loaded structures, hit **File** in the top-level menu and **Save As...** in the next menu. In the resulting *Save File As* dialogue box, specify the name of the file to which you want to write, choose the required format in the *Save as type* entry box, and hit **Save**.
- All currently-loaded molecules are included in the output files. Thus, these formats are useful for transferring assemblies of molecules from the Hermes visualiser into other programs. However, display properties (e.g. atom colours) are not retained.

20 Appendix A: Special Details for SuperStar Users

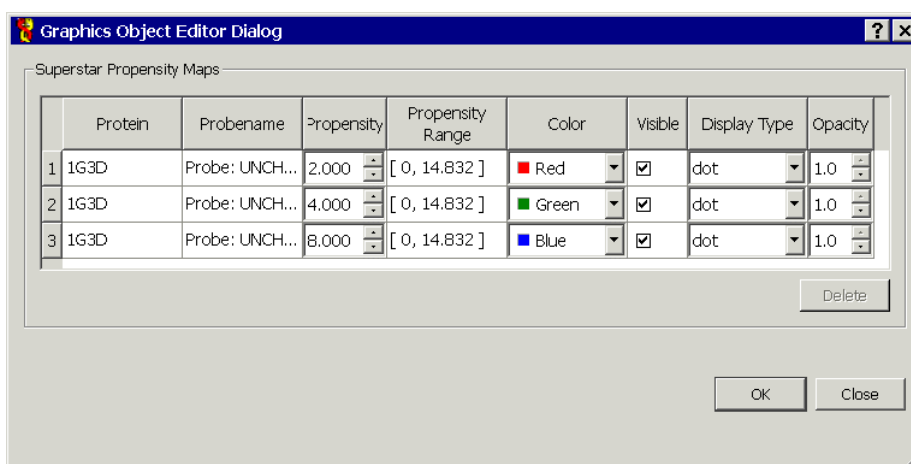
20.1 Toggling and Controlling the Display of SuperStar Generated Maps

20.1.1 Controlling the Display of Propensity, Pharmacophore, Cavity and Connolly Maps

- The appearance of the Hermes visualiser will change depending on which map(s) has been calculated.
- Entries for contour propensity, pharmacophore, cavity and Connolly surface maps will be added to the *Graphics Object Explorer* for each loaded protein as the maps are calculated.



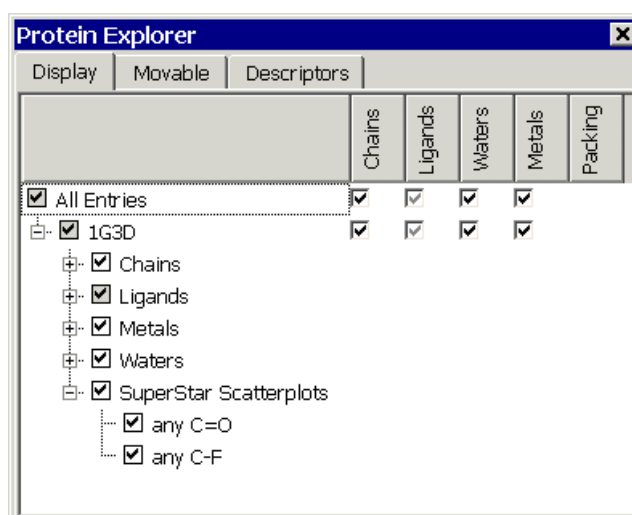
- Whether or not each map is displayed in the 3D view can be controlled by activating or deactivating the tickbox adjacent to the map type.
- All SuperStar generated maps can be switched off by deactivating the tickbox next to *Superstar*.
- The settings for individual maps can be edited by right-clicking on the appropriate map entry in the *SuperStar* tree.
- The options that are presented depend on which map has been right-clicked on. One or more of the following options will be presented:
 - **Edit:** this option will launch a *Graphic Object Editor Dialog* window where a number of parameters can be modified depending on the map selected.



- **Colors:** use this option to globally change the colour the map is displayed in.
- **Surface Type:** this option is specific to propensity contour maps. Use this to change the display of all contours for a particular probe. Note that the settings for individual contours for a particular probe can be modified via **Edit, Display Type**.
- **Delete:** use this option to delete the map from the Graphics Object Explorer and the 3D view.

20.1.2 Controlling the Display of Scatterplots

- The appearance of the Hermes visualiser will change depending on which map(s) has been calculated.
- In the case of scatterplots, an entry is added to the *Protein Explorer* for each map calculated.



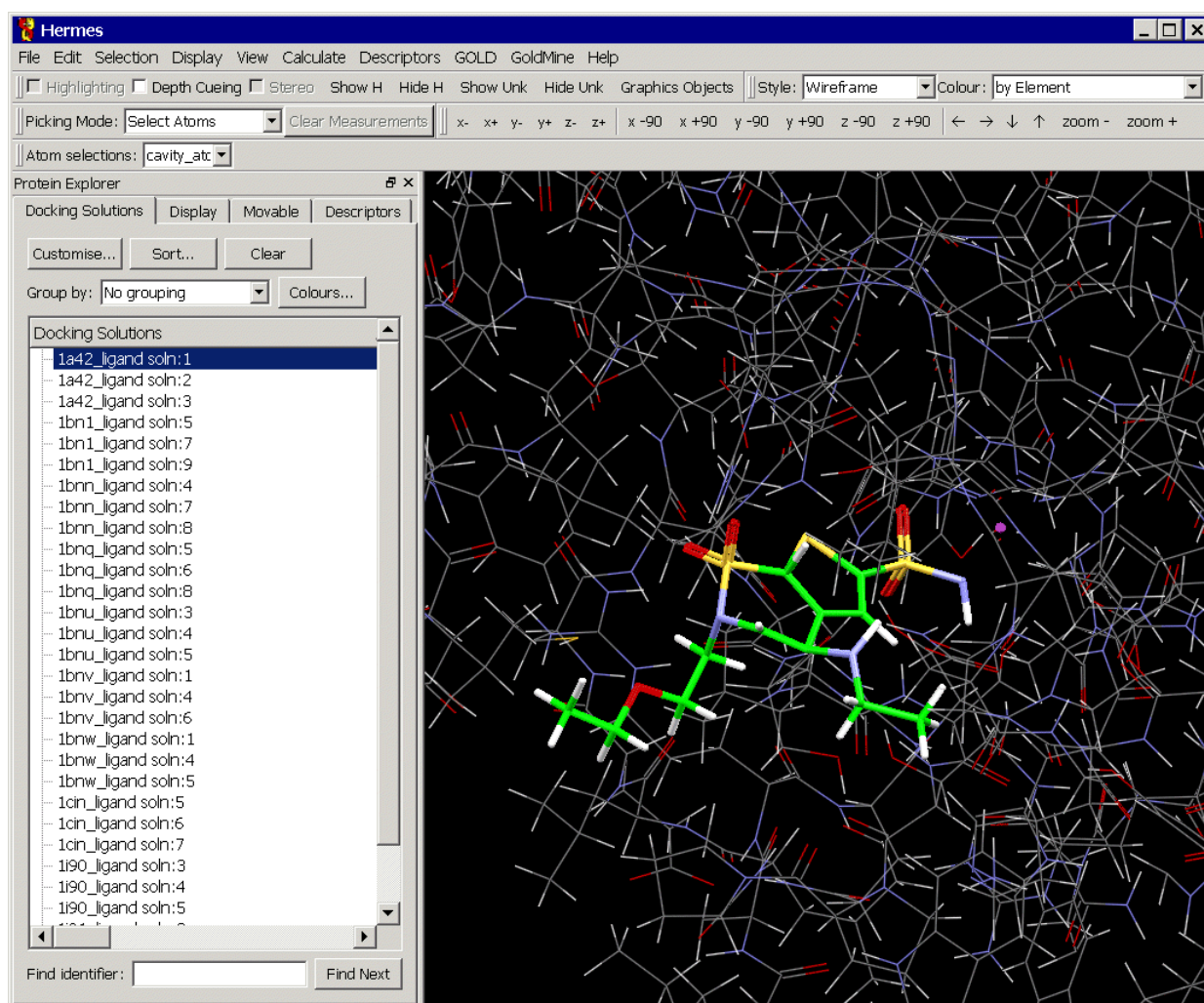
- All scatterplots can be hidden using the tickbox adjacent to the *SuperStar Scatterplots* entry in the tree view. The display of individual scatterplots is controlled using the entries beneath *SuperStar Scatterplots* e.g. *any C=O* above.

- The properties of the scatterplots can be edited by right-clicking on individual *SuperStar Scatterplots* tree entry. Available options are the following:
 - **Styles:** use this option to select the scatterplot display type, e.g. *Wireframe*, *Stick*, *Ball and Stick*, *Spacefill* and whether or not bond types and aromatic rings are displayed.
 - **Colours:** choose to colour by element (default), pick a pre-defined colour or create your own, or customise the carbon colour.
 - **Labels:** use this to label both the scatterplot and input molecule atoms by *Atom Label*, *Element*, *Atom Type*, *Protein Residue*, *Atom File Index* and *Structure*. (see Labelling Atoms, page 37). Use this option also to control the label colour and size.
 - **Select:** use this option to select only the scatterplot atoms. Similar to **Select Only**.
 - **Deselect:** use this option to deselect anything that is selected in the 3D view. Equivalent to clicking in empty space in the 3D view.
 - **Auto Select:** activate this option so that any item subsequently selected in the *Protein Explorer Display* pane will select the corresponding atoms in the 3D visualiser, e.g. clicking on a *SuperStar Scatterplot* entry in the *Protein Explorer* will select all scatterplot items (see Using the Protein Explorer to Make Atom Selections, page 14).

21 Appendix B: Special Details for GOLD Suite Users

21.1 Introduction: Visualising and Refining Selections of Docking Poses

- Docking poses from GOLD docking jobs or from GoldMine databases can be brought into the GoldMine spreadsheet and visualiser for visualisation alongside the associated protein model. Numerical data associated with the solutions, such as the fitness score and its components, are tabulated as columns within the spreadsheet. The spreadsheet data can be sorted and/or colour coded according to selected data columns. Poses can be grouped. Poses may also be manually selected and can then be re-exported with a tailorable number of fields of associated data.
- To read poses in from a GOLD configuration file select **Load GOLD results** from the **File** pull-down in the Hermes top level menu. Browse to the directory in which the relevant GOLD configuration file is situated, highlight on this file and click **Open**.
- The protein file associated with this set of docking results will be displayed.
- In addition the *Protein Explorer* window is brought up and a new tabbed pane displayed within it. This is the *Docking Solutions* pane and it contains a spreadsheet which displays the docking solutions that make up the selection and also descriptor data associated with the selection. The far left-hand column is the *Docking Solutions* column and it contains the full identifier of each solution. It is displayed by default.
- Columns to the right of the *Docking Solutions* column are each headed by the descriptor that they tabulate. Descriptor columns will normally contain those fields that are included in the solution files of the docking job. These will usually be the fitness score and its components.



21.2 Appending and Clearing GOLD Docking Results

- To append a new set of GOLD docking results simply load a new *gold.conf* file.
- To clear all docking solutions without deleting any protein files, click on the **Clear** button in the *Docking Solutions* pane.
- To close all docking solutions and clear the *Protein Explorer* window, click on **Close All Files** in the **File** pull-down on the Hermes top-level menu bar.

21.3 Reading in GOLD Docking Results

- The Hermes visualiser can directly read in the protein and docking poses from a GOLD docking run without first having to create a GoldMine DB. This is useful in cases where only a small number of docking poses have been saved and no filtering of solutions is needed.
- To load the results of a GOLD run, select **Load GOLD results ...** from the **File** pull-down menu. Browse and select the *gold.conf* file relevant to the docking job to be analysed and click on **Open**. Alternatively use Drag-and-Drop as described above, if available.

- The appropriate protein will be displayed. The docking solutions and all the numeric data associated with each solution will be displayed in the *Docking Solutions* window.

21.4 Visualising Individual Poses

- To see the pose for any given docking solution, left-click on the appropriate identifier in the *Docking Solutions* pane.
- If the GoldMine database contains an ensemble docking run, then the protein model from the ensemble that is associated with the docking solution, will be displayed. The *Gold_Ensemble_ID* descriptor can be brought into the spreadsheet to help identify which protein model is being displayed.
- To see multiple poses, keep the **Control** keyboard key down whilst left-clicking the second and subsequent identifiers in the *Docking Solutions* pane.
- To find a pose or poses which have a certain identifier or whose identifiers start with a common sequence, type the sequence into the *Find identifier* box at the bottom of the *Docking Solutions* pane. The spreadsheet will be scrolled so that the first pose starting with that sequence will now occur at the top and will be highlighted in grey.
- To find the next such pose click **Find Next**.

21.5 Picking and Exporting Subsets of Poses

- Once a GoldMine has been created it is possible to create a subset of poses by manually picking from the list within in the *GoldMine* pane. This subset can then be exported in MACCS (SDF), MOL2 or CSV format. It is also possible to export a list of structure names.
Note for GoldMine users: alternatively the subset can be saved as a binary descriptor within an open GoldMine DB. This allows the subset to be used in future Boolean queries.
- To start picking a subset of poses, right-click in an empty space in the *GoldMine* pane. In the pull-down menu that appears, click on **Start Picking**.
- Instructions on how to pick solutions are then provided in the *Picking mode* window. Click on **OK**.
- A new column will appear with the column header *Picks*. All the entries in this column will initially be zero. This column represents a binary or on-off descriptor. The following options are available to pick a subset of solutions:
 - To pick a solution, first left-click on it to select it. Then type ‘a’. A ‘1’ will appear in the *Picks* column. Further solutions may be picked in the same way.
 - To remove an already picked solution, select that solution and type ‘d’.
 - To toggle a value in the *Picks* column from ‘1’ to ‘0’ or vice versa, click on it and type ‘s’.
- When a subset has been created, right-click and select either **Export** or **Export Picks** from the pull-down menu.
 - If you wish to save the subset of poses as a file of structures click on **To a file**. This will bring

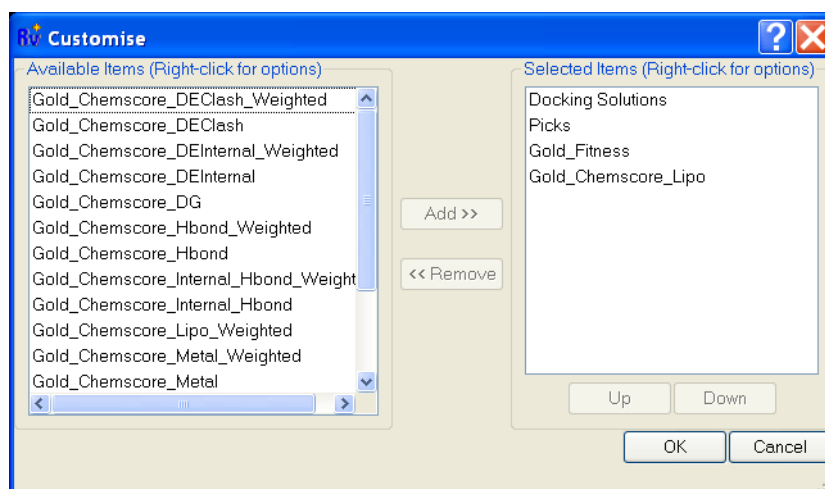
up the *Export Selection* window. Picks and associated descriptor data can be saved as MACCS (SDF) or MOL2 structure files. Alternatively descriptor data can be exported in .csv format. In addition a list of chemical names or identifiers can be generated. This is useful if the intention is to purchase compounds or select from a database for screening. Select the relevant file output from the pull-down menu at the top left corner of the pane. Select those descriptors in the left-hand pane that you wish to include as fields in a MACCS, MOL2 or CSV file and click on **Add**. Enter an appropriate file name in the text box. The **Browse** button will allow you to navigate to the appropriate directory. Click on **Save**.

- If you do not wish to export the picked subset, click on **Not at all**.
- *NOTE to GoldMine users: If you have a GoldMine open and wish to save the Picks column as a binary descriptor in the GoldMine DB, click on **To the GoldMine**. You will then be asked to enter a name for the descriptor. The new descriptor will appear in the Descriptors pane of the GoldMine Controller associated with the appropriate dock set.*
- Right-clicking and selecting **Export picks** from the pull-down menu brings up the same export options described in the previous bullet. This option is available whenever a *Picks* column is displayed.
- *NOTE To GoldMine users: Right-click and select **Import picks** from the pull-down menu to import a picked subset already created for the GoldMine DB currently open. You will be asked to select from the list of picked subsets associated with the open GoldMine DB. Any existing picks in the Picks column will be overwritten.*

21.6 Managing GoldMine Database Data: the GoldMine Pane

21.6.1 Adding, Removing and Ordering Descriptors

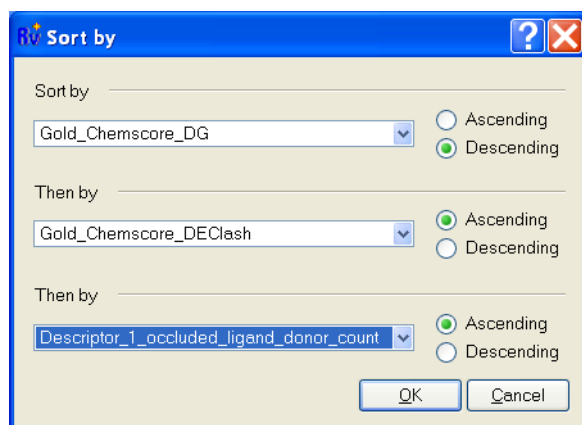
- The **Customise...** button launches a *Customise* window that allows you the option of selecting further descriptors to add to the spreadsheet, or alternatively deleting descriptors that are already in the spreadsheet.



- To add one or more descriptors, select them in the left-hand pane (Left-click and Control select second and subsequent descriptors). Then click on **Add**.
- To remove any of the descriptors select one or more descriptors in the right-hand pane and click on **Remove**.
- To change the order of any of the descriptors in the spreadsheet, first select one or more descriptors in the right-hand pane. The **Up** and **Down** keys at the base of the right-hand pane should now become active. They can be used to change the position of the selected descriptors in the list. This will correspondingly change the ordering of the descriptor columns in the spreadsheet.

21.6.2 Sorting Solutions

- Docking solutions can be sorted using any text or numeric descriptor column. Sort operations can be concatenated.
- A quick way to sort solutions is to click on the column title of the descriptor chosen to sort by. The sort order will be highest to lowest for numerics. Sort order can be reversed by clicking on the column title again.
- An alternative route allows you to sort solutions in an ordered manner. Click on the **Sort...** button. This brings up the *Sort by* window.
- Three different descriptors can be chosen from pull-down menus and placed in the order that they should be sorted. The radio buttons on the right-hand side allow you to select whether each sort should be ascending or descending.



21.6.3 Grouping Solutions

- The default display of docking solutions is in a flat format with every solution displayed in order. However it is also possible to display the solutions in a tree structure where the branches correspond to different groupings.
- Click on the arrow to the right of the *Group by:* box to display the pull-down menu that gives the

grouping options.

- If you select *Ligand only* then all docking solutions for each ligand will be grouped under the identifier for that ligand. Click on the + sign in the box adjacent to the docking solutions identifier, to expand the branch and see data for all the docking solutions relevant to that ligand. This grouping will only be useful if multiple docked poses have been saved for each ligand.
- If you click on a group identifier then a ligand pose from that group will be displayed. A useful trick is to first click on the column title of the descriptor you are most interested in. This sorts the poses within each group according to that descriptor. Then the ligand that is displayed on clicking the group identifier, is the one with the highest score for the sorted descriptor (or alternatively the lowest score, if the column title has been double clicked).
- If you select *Protein only* then all solutions will be grouped under the protein name they belong to. Click on the + sign in the box adjacent to the docking solutions identifier, to expand the branch and see data for all the docking solutions relevant to that set. This grouping will only be useful if results from at least two docking jobs using different protein models have been opened. *NOTE to GoldMine users: the Dock Set name will be used instead of the protein name if the poses being viewed come from an open GoldMine DB. A different dock set may not always contain a set of dockings to a different protein model, it could for instance contain a set of dockings created using a different scoring function. However the grouping will still operate.)*
- If you select *Ligand then Protein* then a tree will be created with the ligand identifier having precedence over the protein identifier.
- If you select *Protein then Ligand* then a tree will be created with the protein identifier taking precedence over the ligand identifier.
- It is possible to also construct a custom grouping. If you select *Custom..* then a *Customise Grouping* window appears. This is very similar to the *Customise....* pane encountered earlier (see Adding, Removing and Ordering Descriptors, page 106). Classifications can be picked in the left-hand pane and transferred to the right-hand pane. The order they appear in the right-hand pane determines the order of precedence of grouping. The **Up** and **down** buttons can be used to change the order of selected classifications. In addition to *Protein* and *Ligand* it is also possible to group by *Pick*. (see Picking and Exporting Subsets of Poses, page 105).

21.6.4 Colouring the Spreadsheet by Temperature Scale

- It is possible to colour descriptor columns in the spreadsheet according to a *Temperature* scale. Thus a graded spectrum of colour is applied with the highest value in the descriptor range coloured red, and the lowest value coloured dark blue. The spectrum is applied over the original range of the descriptor. Colours may be used to rapidly identify high and low points. They may also be useful to display where two descriptors have a correlation.
- Click on the **Colours...** button. The *Customise Colours* window comes up. Selected descriptors can be added to the right-hand window using the **Add** button. The columns representing these descriptors will be coloured according to temperature scale.

- Selected descriptors can be removed from the right hand pane by using the **Remove** button.

21.7 Visualising Saved Protein Subsets Within Hermes

- If you have set up protein subsets (see Making Complex Selections: Defining Subsets of Protein Atoms, page 24) you can highlight the atoms belonging to any one of them by picking the required subset from the *Atom Selections* pull-down menu, which is situated beneath the main Hermes menu.
- Protein files created by GOLD during docking (*gold_protein.mol2* files) contain the set of atoms used by GOLD to define the active site cavity (CAVITY_ATOMS). For any *gold_protein.mol2* files generated by GOLD v3.2 or later and read into Hermes or included in a GoldMine DB, you will have the option in the *Atom selections* window to select a predefined set named **cavity_atoms**.

Note: Any *gold_protein.mol2* generated with previous versions of GOLD can be edited to allow *Cavity Atoms* to be made an *Atom* selection. Add to the end of the second line below @<TRIPOS>MOLECULE five spaces followed by 0 and then five spaces followed by 6.

- Complex selections of protein atoms can be made on the basis of property (e.g. Hydrophobic non-H atoms), atom type, proximity to the ligand etc. (see Making Complex Selections: Defining Subsets of Protein Atoms, page 24). Complex selections are required for the definition of descriptors (see Defining Descriptors, page 75). Complex selections include cavity definitions and other atom lists defined within *gold_protein.mol2* files that are read in or specified within a GoldMine DB (see Visualising Saved Protein Subsets, page 26). These are automatically created.

21.8 Visualising GOLD Fit Points Files

- GOLD outputs a *fit_pts.mol2* file after protein initialisation. This file contains fitting points for the protein active site, i.e. the solvent accessible area of the active site established when the protein is initialised.
- This file can be read into Hermes so that the solvent accessible surface can be visualised either on its own or together with its associated protein. To do this click on **GOLD** and then select **Load GOLD Fitting Points** from the resultant pull down menu. Use the file browser to select the appropriate *fit_pts.mol2* file.

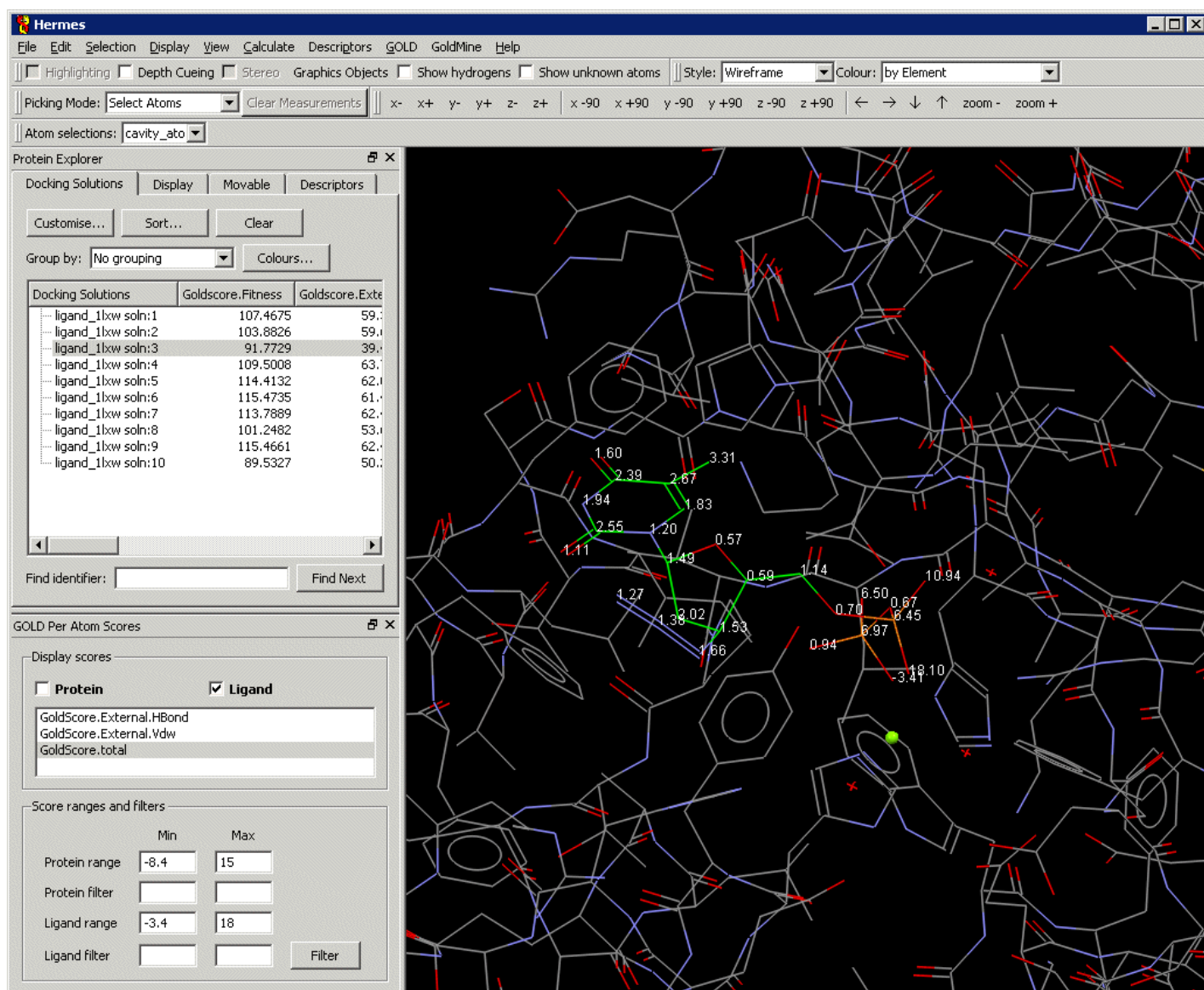
Note: for protein ensembles you will need to ensure the fitting points index number corresponds to that of the associated protein, i.e. that the index number is the same for both the protein and the fitting points file.

21.9 Viewing and Managing Per Atom Scores

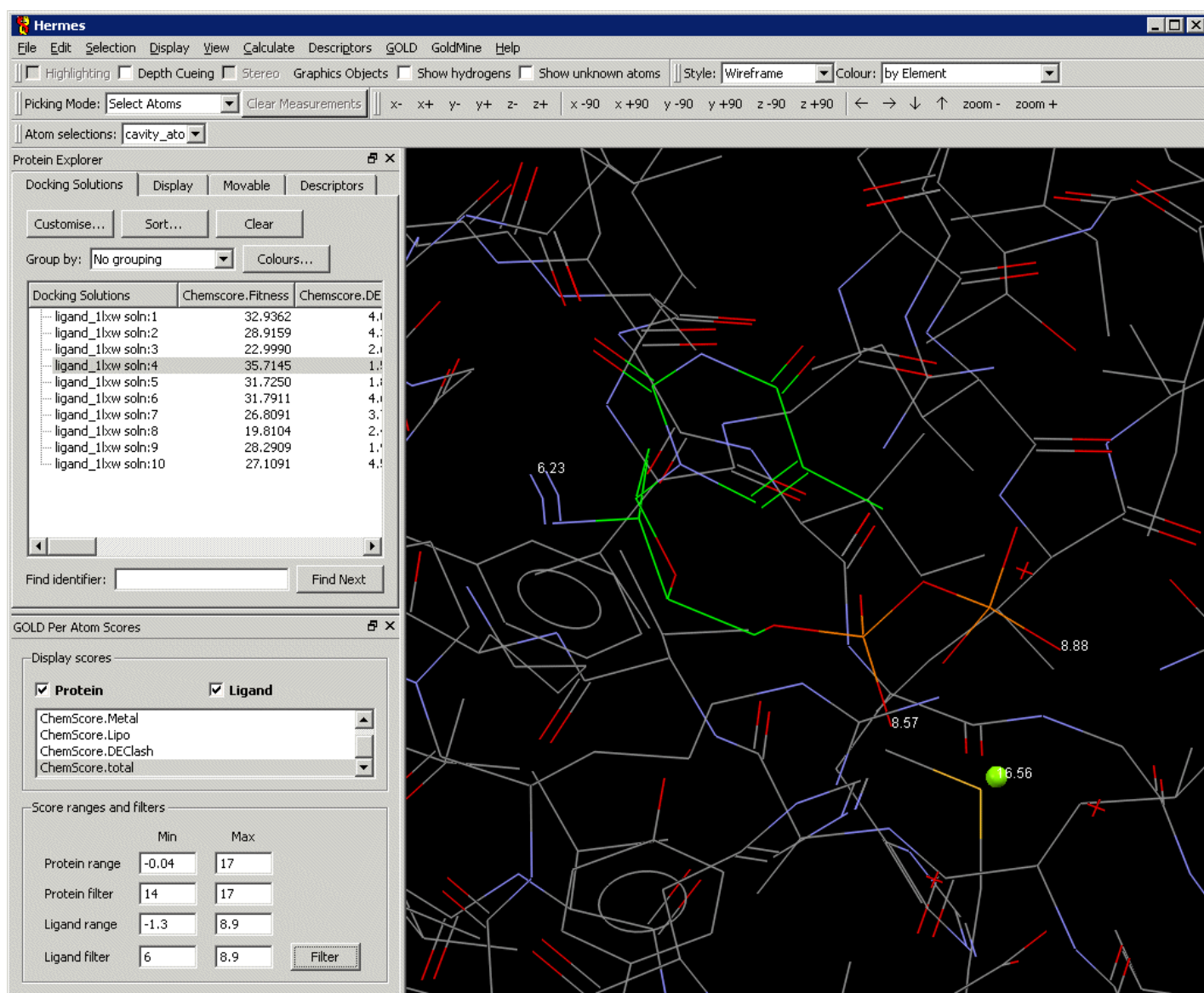
- If per atom scores have been saved during docking, these can be viewed for individual poses by clicking on **GOLD** and then **GOLD Per Atom Scores**.

Note: a pose must first be selected for the *GOLD Per Atom Scores* boxes to be populated and for

the individual atom scores to be displayed in the 3D view.



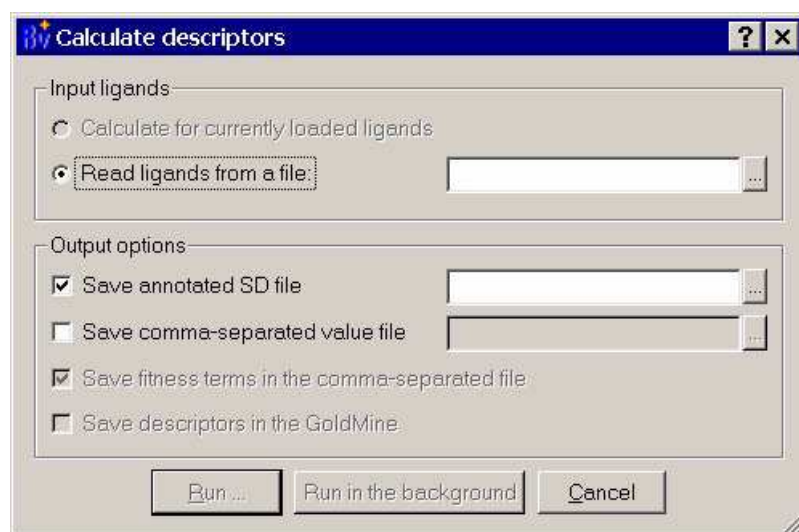
- After a solution has been selected, per atom scores can be displayed for protein or ligand atoms individually or both simultaneously using the **Protein** and **Ligand** tick boxes in the *GOLD Per Atom Scores* dialogue box.
- Each atom's contribution to different components of the chosen scoring function, or the total score, can be displayed by selecting from the options beneath the **Protein** and **Ligand** tick boxes.
- Maximum and minimum atom scores are displayed in the *Protein range* and *Ligand range* boxes.
- To display only those atoms whose scores fall within user-defined values, input the required maximum and minimum scores into the **Protein filter** or **Ligand filter** boxes and select the **Filter** button. The 3D view will update to display only those atoms that satisfy the input values.



21.10 Calculating Descriptors for a Set of Docked Ligands

21.10.1 Specifying the Input to a Descriptor Calculation

- Select **Descriptors** from the top-level menu followed by **Calculate...** from the resulting pull-down menu. This will open the *Calculate descriptors* dialogue:



- If the descriptors are to be calculated from a set of docking solutions from a GoldMine DB, first ensure all the structures for which descriptors need to be calculated have been transferred to the *Docking Solutions* window.
- Turn on the *Calculate for currently loaded ligands* radio button.
- Normally you will want to append the results, as new descriptors, to the GoldMine DB the poses came from. If so, turn on the *Save descriptors in the GoldMine* radio button.
- The results of a descriptor calculation for a large set of docked ligands can be analysed within GoldMine to e.g.
 - to filter out solutions with descriptor values you regard as unfavourable;
 - to sort on descriptor values you feel may correlate with activity;
 - to plot histograms and scattergrams.
 - You may also want to produce plots or perform statistical calculations.
- Instructions on how to use the above features in GoldMine can be found in the GoldMine documentation.
- If calculating from an external file of docking solutions, type the name of the file containing the docking solution(s) into the *Read ligands from a file* entry box (or click on the ... button to select the file via a browser). The ligand file may be in *sdf* or *mol2* format and may contain one or more ligands. Descriptors will be calculated for all docking solutions in the file.
- If results are not to be appended to a GoldMine DB then specify the required output options (see Output Options for Descriptor Calculations, page 112).

21.10.2 Output Options for Descriptor Calculations

- Output may be written out in one of two formats, *sdf* and *csv*. Switch on the *Save annotated SD file* check box for the former or pick *Save comma-separated value file* for the latter.

- If *sdf* format is chosen, the output file will contain all the information (3D coordinates, etc.) that the input file contained, together with additional, tagged data fields containing descriptor names and values, e.g.

```
> <Descriptor_1.molecular.weight>
258.494
```

```
> <Descriptor_1.rotatable.bond.count>
17
```

- If *csv* format is chosen, output will be written to a comma-separated value file. The first line of the file will contain the descriptor names. There will then follow one line for each docking solution containing the descriptor values for that solution, e.g.

```
Entry,Descriptor_1 molecular weight,Descriptor_1 rotatable bond count
LIGAND1, 258.494, 17
```

- Specify the output file name in the entry box adjacent to the output option you have chosen. You may type directly into the entry box or click on the "..." icon to use a file browser. The file extension (*.sdf* or *.csv*) will be added automatically.
- If you have selected *csv* output, switch on the *Save fitness terms in the comma-separated file* box if you want the output file to include the GOLD docking-score terms, i.e. the total GoldScore or ChemScore value for each docking, and its components such as protein-ligand H-bond energy, internal ligand strain energy, etc. These terms will automatically be written out if *sdf* output has been selected.

Note: This option is only relevant when calculating descriptors for GOLD docking solutions.

- Output *csv* files will also include an index column, which will contain the number of each docking solution in the input file (i.e. starting at 1 and finishing at *n*, where there are *n* docking solutions in the file).
- The output files may be large.
- Next, start the calculation (see Running a Descriptor Calculation, page 113).

21.10.3 Running a Descriptor Calculation

- Hit the **Run...** button to calculate the descriptors interactively or **Run in the background** to put the calculation into the background. The calculation may take several hours for a large set of ligands, so the latter option will normally be preferred.
- If you hit **Run...**, a window will be displayed which will monitor the progress of the calculation. You can hit **Interrupt** in this window at any time to abort the calculation (results generated up to that point will not be lost).

- Results will be written to the specified output file (see Output Options for Descriptor Calculations, page 112) which should then be processed by a data-analysis program of your choice (see Using Calculated Descriptors in External Programs, page 91).

21.10.4 Running Descriptor Calculations via the Command Line

- Descriptor calculations can be run directly from the command line. You will require a pre-defined XML file (defined via the Hermes interface) and a gold.conf file. The following command should be used:

```
descriptor_calculator -conf <gold_conf_file> -xml <xml_file> -
csvout <csv_output_filename> -sdout <sd_output_filename>
```

Alternatively is possible to specify a ligand filename in *nogui* mode, using the command:

```
descriptor_calculator -ligands <ligand_file_name> -xml
<xml_file> -csvout <csv_output_filename> -sdout
<sd_output_filename>
```

Note: You do not have to specify both *sdout* and *csvout* output if you do not want to.

21.11 Saving Docking Files from Hermes

- When a GoldMine is loaded and structures are being viewed, a number of output options are available from the **File** menu:
 - **Save As:** allows you to save out the protein file in **.pdb*, **.mol2*, **.reliarch*, **.bmp*, **.jpg*, **.png* or **.pov* format.
 - **Export:** use this option to export all ligands currently loaded in the 3D view, including user-defined tags. Ligands can be output in **.sdf*, **.mol2*, **.csv*, **.txt*, or **.html* format.
 - **Export Complex:** use this option to output the protein-ligand complex currently on display. Choose from **.pdb* and **.mol2* output formats.

22 Appendix C: Special Details for Relibase+ Users

22.1 Recommended Settings for Hermes when using Relibase+

- When using Hermes with Relibase+ we recommend the following settings used, both of which can be enabled or disabled via the **Display** main menu option:
 - **Overwrite on load** should be switched on (i.e. with a tick adjacent to the text): this means that whenever a new structure is launched in Hermes it will overwrite what is currently on display, completely refreshing the Hermes 3D view with the new structure.
 - **Allow multiple instances** should be switched off (i.e. with no tick adjacent to the text): this means that only one Hermes client will be used during a Relibase+ session.

23 Appendix D: Calculating Descriptors: Computational Details

23.1 Acceptor Atoms

An acceptor atom is any atom capable of accepting a hydrogen bond. This includes almost all oxygen and nitrogen atoms, but with the exceptions of trigonal planar (R_3N) and quaternary (R_4N) nitrogens, since these have no free lone pairs. Oxygen atoms in conjugated environments (e.g. furan oxygen) and the central nitrogen of an azide group are counted as acceptors, though this is questionable.

23.2 Atom Types

Some options in the Relibase+ visualiser allow specification of Sybyl atom types, devised by Tripos Inc. (<http://www.tripos.com>). Amongst the most common types are:

H	hydrogen
C.1	<i>sp</i> carbon
C.2	<i>sp</i> ² carbon
C.3	tetrahedral carbon
C.ar	aromatic carbon
N.1	<i>sp</i> nitrogen
N.2	two-coordinate, non-linear nitrogen
N.3	pyramidal trigonal nitrogen
N.pl3	planar, trigonal nitrogen
N.am	amide nitrogen
N.ar	aromatic nitrogen
O.2	carbonyl oxygen
O.3	ether oxygen
O.co2	carboxylate/phosphate oxygen
S.2	thiocarbonyl sulphur
S.3	thioether sulphur
S.o	sulphoxide sulphur
S.o2	sulphone sulphur
P.3	phosphate phosphorus

Halogens	standard element symbols e.g. F, Cl, Br, I
Metals	standard element symbols, e.g. Zn, Fe, Ca

23.3Buried Atoms

An atom is counted as buried if no part of it is solvent accessible (see Solvent Accessibility, page 119). It does not matter what types of atoms render the atom inaccessible, i.e. whether the atom is buried by protein or ligand atoms or a combination of both.

23.4Donatable (or Polar) Hydrogens

A donatable or polar hydrogen is any hydrogen that can be donated in a hydrogen bond. This includes any oxygen-, nitrogen- or sulphur-bound H atom.

23.5Hydrogen Bonds

By default, a protein-ligand hydrogen bond is present if:

- it involves a recognised polar hydrogen (see Donatable (or Polar) Hydrogens, page 118) and a recognised acceptor atom (see Acceptor Atoms, page 117); and
- the H...acceptor distance is less than the sum of van der Waals radii and the donor-H...acceptor angle is greater than 90.

These criteria can be customised when setting up H-bond descriptors (see Customising the Geometrical Definition of a Hydrogen Bond, page 83). Intramolecular hydrogen bonds, either in the protein or the ligand, are not recognised.

23.6Hydrophobic Atoms

The definition of hydrophobic atoms includes almost all types of carbon atom (but not cyanide or carbonyl carbon), sp^3 hybridised sulphur, non-ionised chlorine, bromine and iodine, and any hydrogen atom that is covalently bonded to a hydrophobic atom (effectively, C-H and S-H). Note that S-H is also counted as a donatable hydrogen.

23.7Number of Buried Hydrogens/Acceptors Not Forming an H-Bond

The following descriptors:

- *Number of buried donatable ligand hydrogen atoms not forming an H-bond*
- *Number of buried ligand acceptors not forming an H-bond*
- *Number of buried donatable protein hydrogens not forming an H-bond*
- *Number of buried protein acceptors not forming an H-bond*

are counts of "occluded" hydrogen-bonding atoms, i.e. atoms that are inherently capable of

participating in hydrogen bonds but are prevented from doing so because they are buried by non-H-bonding atoms. For example, a histidine-ring NH would be counted as occluded if it were solvent accessible before docking but was buried by ligand carbon atoms after docking. A limitation of these descriptors is that no allowance is made for intramolecular hydrogen bonding (i.e. in the example just quoted, the NH would still be counted as occluded even if it were H-bonded to a neighbouring protein atom).

23.8 Polar, non-Hydrogen-Bonding Atoms

Polar atoms in descriptors such as *Number of buried polar atoms in the ligand* are atoms that have some polar character but cannot participate in hydrogen bonds. The definition of these includes any nitrogen that is not counted as an H-bond acceptor (notably planar-trigonal and quaternary N) and fluorine, sulphur and phosphorus.

23.9 Rotatable Bonds

A ligand bond is considered rotatable if it is single, acyclic and not to a terminal atom. This therefore includes, e.g., bonds to methyl groups but not to chloro substituents. It also includes bonds which, although single and acyclic, have highly restricted rotation, e.g. ester linkages. Finally, it incorrectly include bonds to linear groups, e.g. the bond between the methyl and cyanide carbons in $\text{CH}_3\text{-CN}$.

23.10 Solvent Accessibility

An atom is counted as being solvent accessible if any part of it is exposed to solvent. Solvent accessibility is measured assuming a default solvent radius of $1.4\text{\AA}$, however this value can be altered.

23.11 Solvent Inaccessible Ligand Surface Area

This is a measure of how much of the ligand surface area is desolvated upon docking, calculated assuming a solvent molecule radius of $1.4\text{\AA}$ (this value is the default value which can be altered). The surface area of the undocked ligand *in the same conformation* is also written out. Values can be output in units of $\text{\AA}^2$ or as percentages. If percentage quantities have been requested, the undocked value will always be zero.

23.12 Sphere Accessible Volume

Two values are output, before and after docking. The former (undocked) value measures how much of the sphere volume is vacant (i.e. unoccupied by protein atoms) before the ligand is docked. The latter measures how much sphere volume remains vacant (i.e. unoccupied by either protein or ligand atoms) after docking. Volumes may be output in $\text{\AA}^3$ or as percentages. If percentage quantities have been requested, the undocked value will always be 100 (since, by definition, 100% of the sphere volume not occupied by protein atoms must be vacant before the ligand is docked). If absolute ($\text{\AA}^3$) values have been requested, the undocked value may vary a little from one docking to another

because GOLD may move protein polar H-atoms during docking, therefore altering the amount of sphere volume that is occupied by protein atoms.

24 Appendix E: Using Descriptors Tutorial

This tutorial has three parts. Parts 1 and 2 must be done in the same Hermes session (i.e. you cannot close the program down part-way through). Once Part 2 is completed, you can, if you wish, close Hermes down and do Part 3 later.

The Example Protein

The tutorial is based on a structure (PDB code 2ACK) of the serine protease acetylcholinesterase (AChE) from *Torpedo californica*. AChE is a target for anti-Alzheimer's drugs and the site of action of several insecticides. A key aim of Hermes is to allow you to set up customised descriptors capturing information that is known to be important for the particular protein being studied. So, we will begin by summarising some key features of the active site and its known structure-activity:

- *The catalytic triad of Ser200, His440 and Glu327*. Hydrolysis of acetylcholine, the reaction catalysed by AChE, involves nucleophilic attack by the hydroxyl oxygen of Ser200 with transfer of the serine hydroxyl hydrogen to NE2 of His440.
- *The oxyanion hole*, formed by the main-chain NH groups of Gly118, Gly119 and Ala201. These groups hydrogen bond to the anionic oxygen of the transition state, thereby stabilising it.
- *Trp84*, which helps bind the substrate by forming a cation - π interaction with the acetylcholine NMe_3^+ group. Several known AChE inhibitors have quaternary ammonium groups or electron deficient aromatic rings that bind close to Trp84 and presumably make an important contribution to binding energy.
- Several other hydrophobic residues, notably *Phe288, Phe290, Phe330 and Phe331*, which form hydrophobic interactions with several known inhibitors.
- *Trp279*, which is a peripheral binding site for electron deficient, hydrophobic groups such as NMe_3^+ . For example, the inhibitor decamethonium has two such groups, one of which binds to Trp84, the other to Trp279.

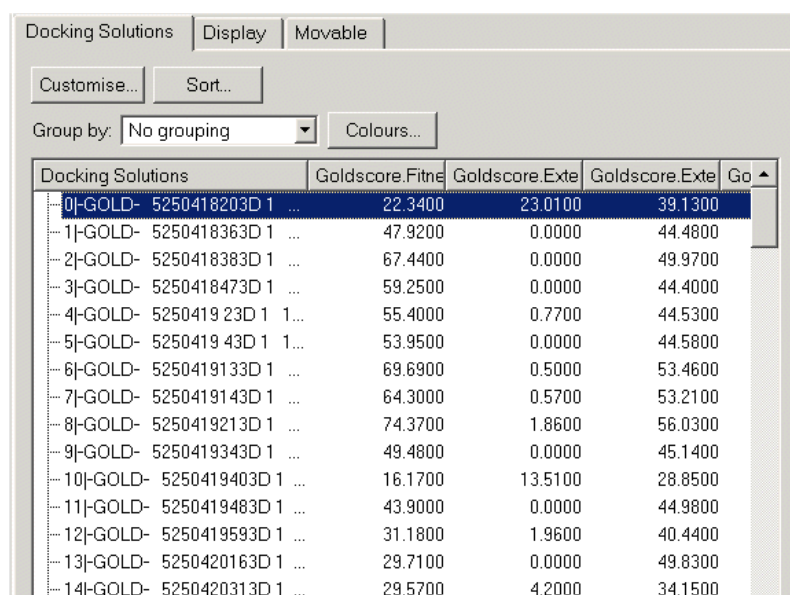
The Example Ligands

We have used GOLD to dock into AChE two known inhibitors and 285 random molecules from the MDPI database (<http://www.mdpi.org>). The GoldScore scoring function was used and the top-scoring docking solution for each molecule kept. We will read these results into Hermes, browse through one or two, set up some protein subsets and define and calculate some descriptors.

24.1Part 1: Visualisation, Browsing, and Creating Protein Subsets

24.1.1 Loading a GOLD Run

- Open Hermes.
- Select **File** from the top-level menu, followed by **Load GOLD results**. In the resulting file dialogue, specify the name of the GOLD .conf file from the docking run (viz. *gold_gs_top1.conf*) and hit **Open** to load the docking results.
 - If you are on a Windows PC, this file will be in the folder `hermes\examples\` in the folder where Hermes is installed, this is most likely to be:
`C:\Program Files\CCDC\gold suite\Hermes\examples`
 - On Unix it will be in the directory `hermes/examples/` in the directory where Hermes is installed, for example:
`<INSTALLDIR>/GOLD_Suite/`
- The *Docking Solutions* list box on the left of the main Hermes window now contains the identifiers of the docked molecules together with the docking score and the components that make up the scoring function. For example, the first molecule has the identifier *5250418203D* (this has been read from the molecule's sd file) and a GoldScore value of *22.34*:



Docking Solutions	Goldscore.Fitne	Goldscore.Exte	Goldscore.Exte	Go
0 -GOLD- 5250418203D 1 ...	22.3400	23.0100	39.1300	
1 -GOLD- 5250418363D 1 ...	47.9200	0.0000	44.4800	
2 -GOLD- 5250418383D 1 ...	67.4400	0.0000	49.9700	
3 -GOLD- 5250418473D 1 ...	59.2500	0.0000	44.4000	
4 -GOLD- 5250419 23D 1 1...	55.4000	0.7700	44.5300	
5 -GOLD- 5250419 43D 1 1...	53.9500	0.0000	44.5800	
6 -GOLD- 5250419133D 1 ...	69.6900	0.5000	53.4600	
7 -GOLD- 5250419143D 1 ...	64.3000	0.5700	53.2100	
8 -GOLD- 5250419213D 1 ...	74.3700	1.8600	56.0300	
9 -GOLD- 5250419343D 1 ...	49.4800	0.0000	45.1400	
10 -GOLD- 5250419403D 1 ...	16.1700	13.5100	28.8500	
11 -GOLD- 5250419483D 1 ...	43.9000	0.0000	44.9800	
12 -GOLD- 5250419593D 1 ...	31.1800	1.9600	40.4400	
13 -GOLD- 5250420163D 1 ...	29.7100	0.0000	49.8300	
14 -GOLD- 5250420313D 1 ...	29.5700	4.2000	34.1500	

- By default, only the protein is displayed in the 3D visualiser. Ligands can be loaded by selecting their identifier. The carbon atoms of the protein are shown in grey, those of the ligands in green. (Docked molecules are referred to as *ligands* throughout Hermes, irrespective of whether they bind in reality.)

24.1.2 Using the Visualiser

- You can rotate (or zoom) the 3D image by moving the mouse while keeping the left (or right) button depressed. Translation is achieved by moving the mouse while depressing both the left mouse button and the *Control* key.

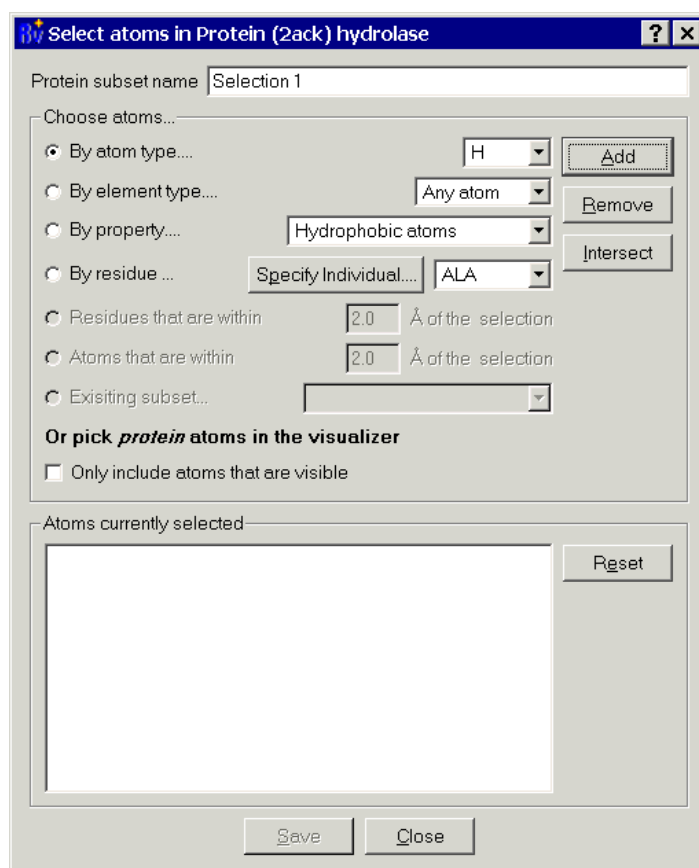
- A wide range of visualisation options can be accessed by clicking the right-hand mouse button. The menu that you get depends on whether the cursor is positioned on an atom, a bond, or neither. For example, right-click anywhere in the visualiser background area (not on an atom or bond), and then, in the resulting menu, click on **Styles** then **Display Aromatic Rings** to toggle off the display of circles within aromatic rings.
- Hide the hydrogen atoms by hitting the **Hide H** button at the top of the Hermes interface.
Note: this option will only switch off the H atoms of the protein and ligand(s) currently loaded.

24.1.3 Browsing Docked Ligands

- Display the fifth ligand in the visualiser by clicking on the fifth line (*4|-GOLD- 525041923D*) in the *Docking Solutions* list box.
- Display short protein-ligand clashes: click on **View** in the top level menu then pick **Contacts**. If contact control window does not appear in a convenient place, it can be docked, i.e. dragged and dropped into the *Docking Solutions* area.
- Click on the *Short Contacts* box display short protein-ligand clashes. Contacts are colour-coded (from green to red) depending on how short they are relative to the sum (**S**) of the van der Waals radii of the atoms involved. Only contacts that are at least 0.5Å shorter than **S** are shown. (Also, to reduce clutter, only H...H and non-H...non-H short contacts are shown, not H...non-H. If a short H...non-H contact is present, there will almost certainly be one of the other two types of bumps too, so this does not lose any information of importance.)
- Turn off the display of short contacts and activate the H-bonds box. This displays the single H-bond that occurs between the protein and this docked ligand. It has a relatively poor geometry, being well out of the plane of the ligand carbonyl acceptor group; Hermes is fairly forgiving in what it classes as a hydrogen bond.
- Load multiple ligands by holding down the *Control* keyboard key and selecting two or three items in the **Docking Solutions** list box. All of the ligands that you have clicked on are now loaded into the 3D display area (loaded ligands are shown in dark blue in the list box). The short-contact and H-bond displays refer to the last ligand that you loaded.
Note: a continuous selection of ligands can be loaded by selecting one ligand, holding down the *Shift* keyboard key, then selecting the last ligand in the continuous list. Alternatively select the first ligand with the mouse cursor and keep the left mouse button depressed whilst moving down the solutions list.
- Now go back to displaying the first ligand only by clicking on the first item (*1-GOLD- 5250418203*) of the *Docking Solutions* list. Turn off the display of short contacts and H-bonds.

24.1.4 Creating and Displaying a Simple Protein Subset

- Protein subsets are frequently required for creating descriptors.
- Pick **Selection** from the top-level menu, followed by **Define Complex Selection**. This opens a *Select atoms in Protein (2ack) hydrolase* window:



- We will create a complex selection containing some of the key catalytic and binding residues, so type a suitable name, *key residues*, into the *Protein subset name* box.
- Complex selections can be created in many ways but we will do it here by specifying individual residues by their number.
- Click on the *By residue...* option and then hit the **Specify Individual...** button. The resulting window lists all the residues. Switch on the check-boxes for *Trp84*, *Gly118*, *Gly119*, *Ser200*, *Ala201*, *Trp279*, *Phe288*, *Phe290*, *Glu327*, *Phe330*, *Phe331* and *His440*. Then hit the **Add** button to add the atoms of these residues into the subset we are building. Hit **Close** to close the *Select individual residues or atoms* window.
- These are all the atoms we want in this subset, so hit **Save** and then **Close** in the *Define a new protein subset* window. The *key residues* subset is now created.
- We can highlight the subset in the 3D display by clicking on the down-arrow icon next to the box labelled *Atom selections* (above the 3D window) and selecting *key residues* from the resulting list. The subset is now selected, i.e. all atoms comprising the subset have yellow spheres on them.
- Which atoms are displayed can now be controlled by right-clicking in the 3D display and selecting one of the options under the **Show/Hide** menu.
- In viewing and moving this display, you may find it helpful to right-click on a central atom of

the ligand and select **Rotation Centre** in the resulting menu, followed by **Set Rotation Centre** in the next menu.

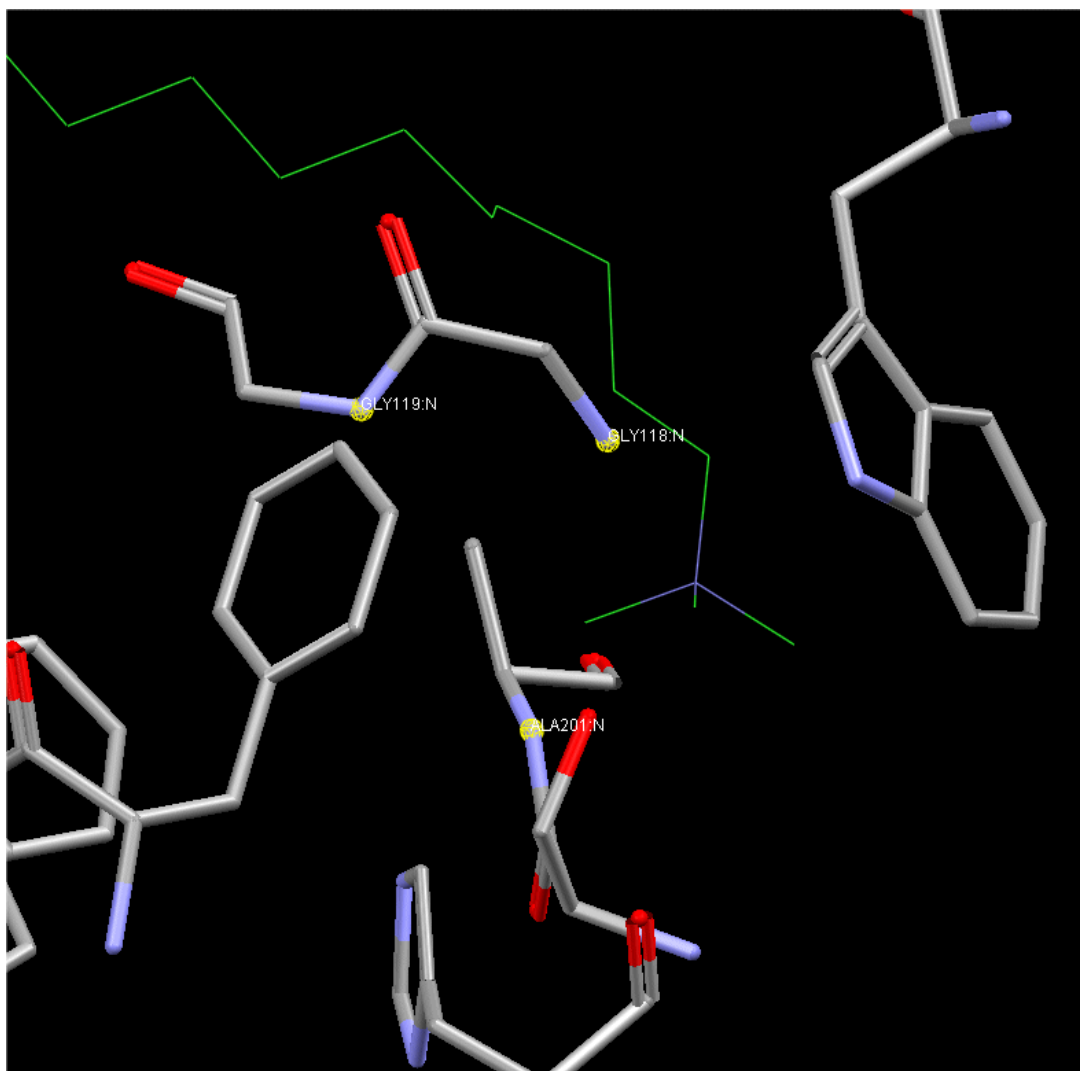
24.1.5 Creating a Larger Complex Selection

- We will now create a subset containing all the key residues together with all other residues near the currently-displayed ligand.
- You must first select the ligand so right click on any ligand atom and pick **Selection** and **Select Molecule** from the resultant pull-down menus.
- Now click on **Selection** in the top-level menu, followed by **Define complex selection**. You will notice the selected ligand atoms are listed in the *Atoms Currently Selected* window. Type the name *active site + other key* into the *Protein subset name* entry box.
- Select the *Residues that are within* option and type in *5.0* into the entry box.
- Hit **Add** to add into this new subset all residues that have at least one atom within 5Å of the ligand currently being displayed. These atoms will be added to the *Atoms currently selected* window and will become selected in the visualiser.
- Now select the *Existing subset* option and ensure that the box alongside is set to the *key residues* subset (it should be, as it is the only subset we have created so far).
- Hit **Add**. This adds to our new subset any atom in the *key residues* subset that is not already in our new subset.
- Now hit **Save** to complete creation of the new subset.
- Our use of **Add** performed a logical OR. Had we used **Intersect**, we would have applied an AND operation (i.e. the subset would have contained all the atoms of key residues within 5Å of the ligand). Use of the various options for choosing atoms, combined with the logical-operation buttons, gives great control over the contents of subsets.
- **Close** the *Define a new protein subset* dialogue and use the *Atom selections* box to the display the new subset. Hide the rest of the protein by right clicking and selecting **Show/Hide** then **Show only**.

24.1.6 Creating a Protein Subset by Directly Picking Atom

- For use in one of the descriptors, we will now set up a subset containing the hydrogen-bond donor groups of the oxyanion hole.
- Open the *Select atoms in Protein (2ack) hydrolase* dialogue again and type in the subset name *oxyanion hole*. If there are still atoms listed from the previous selection, hit **Reset** next to the *Atoms currently selected* section of the window.
- In the main Hermes dialogue use the *Atom selections* box to display the *key residues* subset and turn off the display of protein hydrogens (this is just to simplify the image to make the next step easier). You can also show only those residues that make up the *key residues* subset using the **Show/Hide** method described above.

- Click on the three backbone nitrogen atoms of the residues involved in the oxyanion hole (Gly118, Gly119, Ala201; you can right-click in the display-area background and select **Labels** ... **Show labels** if you struggle to find the residues).



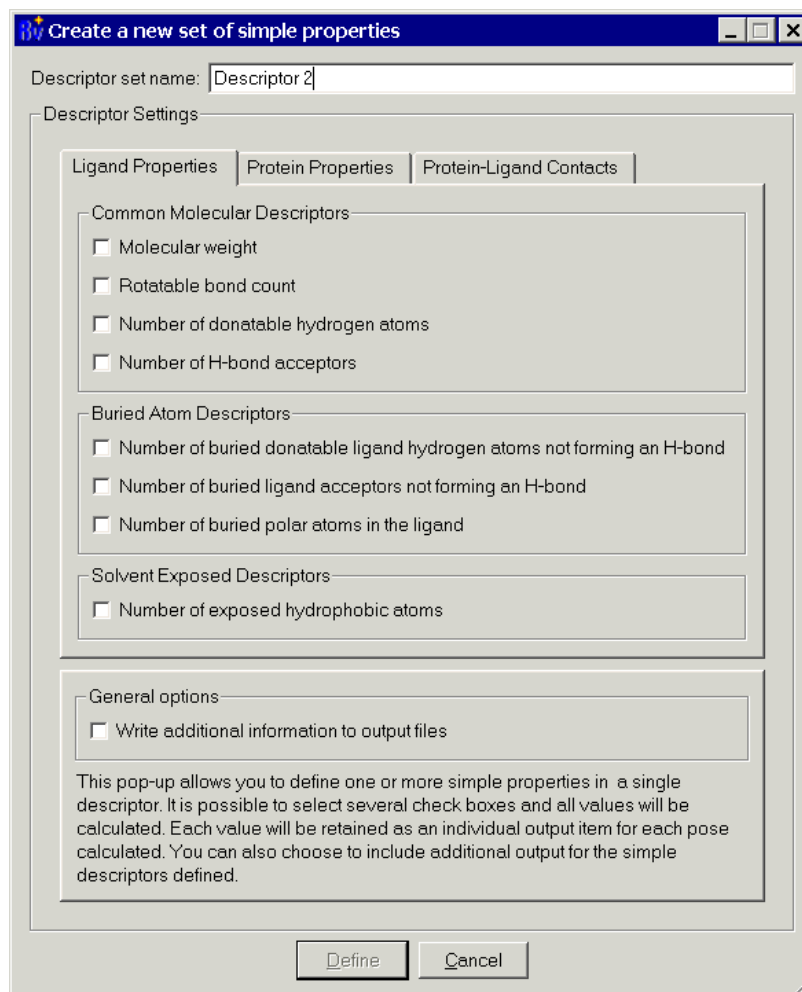
- At this point, the three N atoms should be selected (marked with small yellow spheres) in the display area and the white box in the bottom part of the *Select atoms in Protein (2ack) hydrolase* dialogue will contain (only) the atom labels *GLY118*, *GLY119* and *ALA201*. If it does not (e.g. if you hit the wrong atoms, or you accidentally had some atoms selected before you opened the dialogue), hit **Reset** and try again.
- Hit **Save** then close the *Select atoms in Protein (2ack) hydrolase* dialogue.

Without closing Hermes, proceed to part 2 of the tutorial.

24.2Part 2: Defining and Saving a Set of Descriptors

24.2.1 Defining a Simple Ligand Descriptor

- Although raw docking scoring-function values are often poorly correlated with affinity, the situation can sometimes be improved by normalising the score by ligand molecular weight.
- Molecular weight is one of a number of simple pre-defined descriptors available in Hermes, but it will not be calculated unless you ask for it.
- Click on **Descriptors** in the top-level menu, then **Define** in the next menu, followed by **Add Simple Descriptor**. This will open the *Create a new set of simple properties* dialogue:



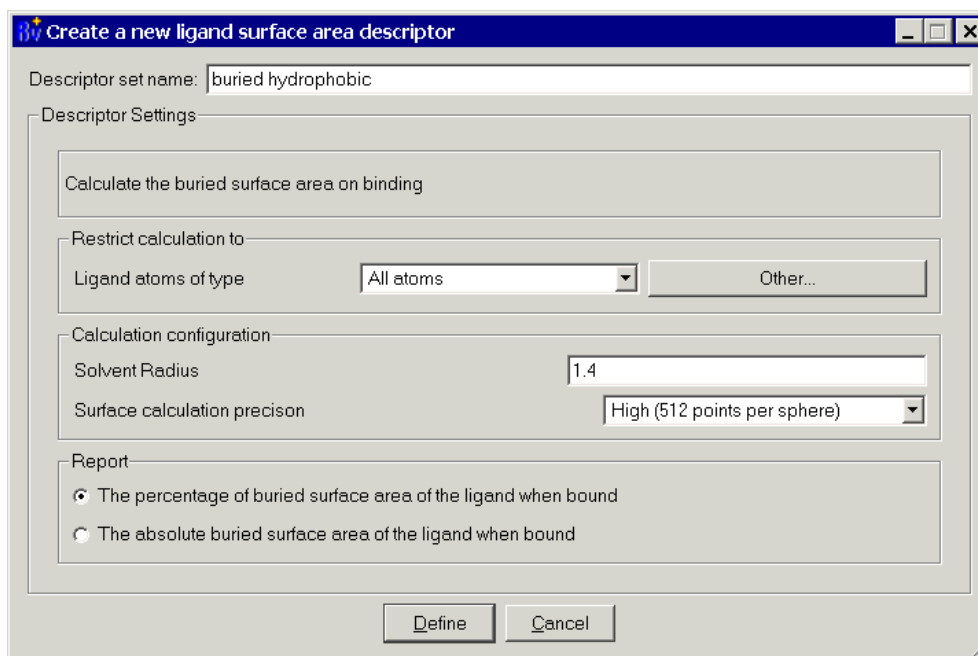
- Type a suitable name, *molec wt*, into the *Descriptor Name* entry box, click on the check box labelled *Molecular weight*, then hit **Define**.
- Go to **Descriptors** again and select **Descriptor Browser**. A new tab called **Descriptors** will be added to the *Docking Solutions* list menu.
- The descriptor we have just defined, *molec wt*, is listed here. As we will see later, once we have defined the other descriptors we want, all defined descriptors can be saved to a descriptor configuration file (.xml format) for future use and/or can be calculated for our set of docking solutions.

24.2.2 Defining Other Simple Descriptors

- Once again, click on **Descriptors** in the top-level menu, then **Define** followed by **Add Simple Descriptor**. Click on the *Protein-Ligand Contacts* tab.
- Type in the name *bump/H-bond count* into the *Descriptor Name* entry box.
- Switch on the check boxes labelled *Number of ligand atoms forming H-bonds* and *Number of ligand atoms forming bad clashes*. More specifically, these will be a count of the number of ligand atoms forming H-bonds *to the protein*, and the number forming (non H-bonding) contacts to the protein that are at least 0.5Å shorter than the sum of van der Waals radii.
- Hit **Define**. The newly defined descriptor is added to the *Descriptor Browser* on the left of the main Hermes window. This descriptor, of course, will result in two separate numbers for each docking solution.

24.2.3 Defining a Ligand Surface Area Descriptor

- We will define a descriptor that represents how much of the hydrophobic portion of the docked ligand surface is solvent inaccessible.
- Click on **Descriptors** in the top-level menu, then **Define** in the next menu, followed by **Add Ligand Surface Area Descriptor**. This will open the *Create a new ligand surface area descriptor* dialogue:

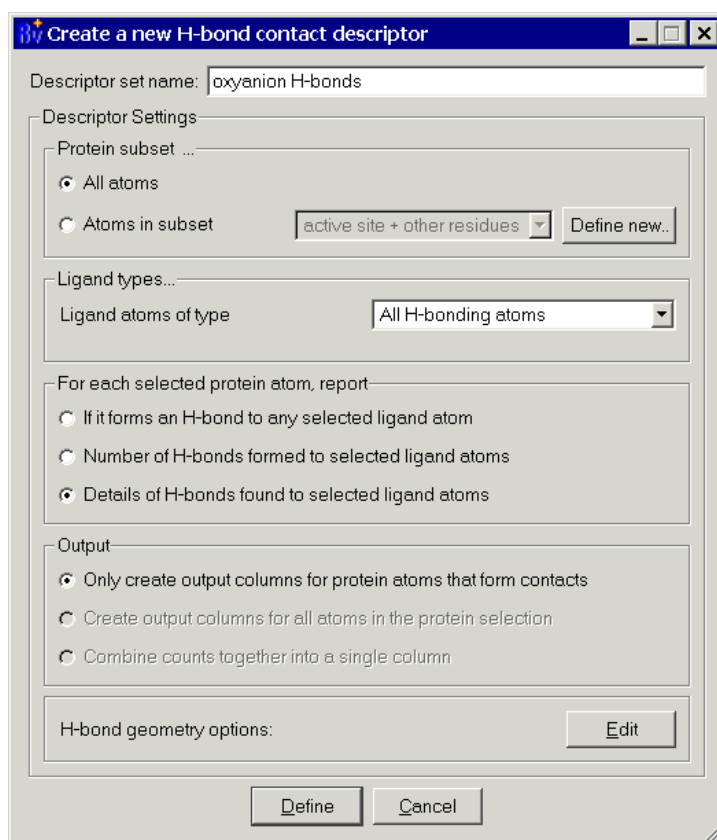


- Type the name *buried hydrophobic* into the *Descriptor Set Name* entry box.
- Hit the down-arrow icon next to the box labelled *Ligand atoms of type* and select *All hydrophobic atoms* from the resulting pull-down menu.

- The *Solvent Radius* value and the settings for *Surface calculation precision* can be altered to configure the speed at which this descriptor is calculated at, however these will be kept at the default settings.
- We have a choice of calculating the absolute amount (in Å²) of hydrophobic ligand surface area that is solvent inaccessible after docking, or the percentage amount. Choose the percentage.
- Hit **Define** to complete the descriptor definition.

24.2.4 Defining a Customised Protein-Ligand H-Bonds Descriptor

- It might be important to know whether any of the three N-H groups of the oxyanion hole are H-bonded to the ligand. You will recall that we have set up a protein subset, *oxyanion hole*, containing these atoms.
- Click on **Descriptors** in the top-level menu, then **Define** in the next menu, followed by **Add H-bond descriptor**. This will open the *Create a new H-bond descriptor* dialogue:



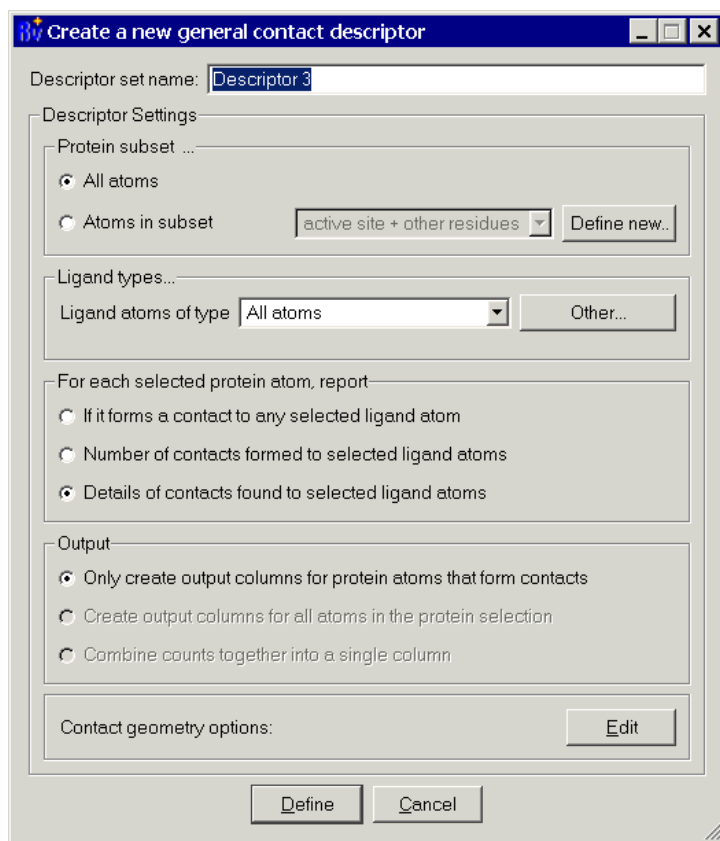
- Set the descriptor name to *oxyanion H-bonds*.
- Select the option labelled *Atoms in subset* and set the box alongside to the *oxyanion hole* subset (by clicking on the down-arrow icon and selecting from the resultant pull-down menu). This will result in a search for H-bonds between these specific protein atoms and the ligand.
- Accept the default options in the remainder of the dialogue box. This means that the descriptor

will include H-bonds between the three protein N-H groups and any relevant type of ligand atom; will count the number of H-bonds that each N-H group forms; and will accept Hermes's default geometric criteria for hydrogen bonds.

- Hit **Define** to complete the descriptor definition.

24.2.5 Defining a Customised Protein-Ligand Contacts Descriptor

- The simple count of ligand atoms forming bad clashes to the protein (see Defining Other Simple Descriptors, page 128) has a hard-coded definition of what constitutes a bad clash. It would be useful to identify docked ligands that form *very* bad clashes, since these ligands could reasonably be dismissed.
- Click on **Descriptors** in the top-level menu, then **Define** in the next menu, followed by **Add General Contact Descriptor**. This will open the *Create a new general contact descriptor* dialogue:



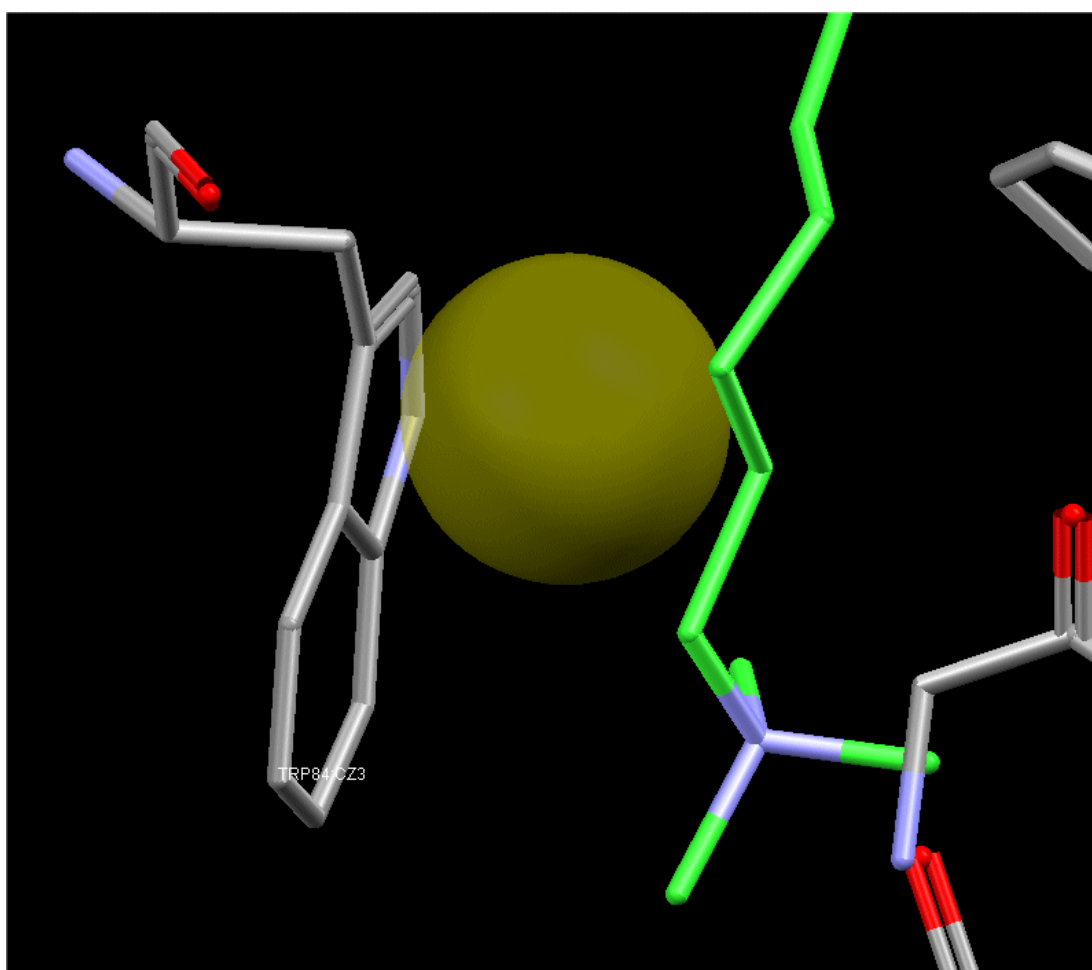
- Set the descriptor name to *very bad bumps*.
- Select the *Atoms in subset* option and set the box alongside to the *active site + other key* subset. Leave the next (i.e. ligand atom type) box at the default *All atoms* setting.
- Select the *Number of contacts formed to selected ligand atoms* option. Since we left the ligand atom-type selection at the default setting, "any selected ligand atom" means, in this case, all

ligand atoms. Thus, the descriptor will search for close contacts between all ligand atoms and protein atoms in active site and other key residues.

- Now we need to specify that only *very* short contacts are required. Hit the **Edit** button to open the *Specify contact definition* dialogue. By default, the *Vdw corrected distance* option is switched on; keep this setting to ensure that contact distances are measured relative to the sum of the van der Waals radii of the interacting atoms. Then set the *Maximum corrected distance* box to *-1.0*. This will ensure that the descriptor only reports contacts that are at least 1.0Å shorter than the sum of van der Waals radii.
- Hit **OK** to close the *Specify Contact Definition* window then **Define** to complete the descriptor definition.

24.2.6 Defining a Sphere Descriptor for Probing Occupied Volume

- Since Trp84 is known to be a key binding residue, it will be interesting to determine which ligands occupy space near to this residue. We can do this by placing one or more spheres in the required region and measuring how much of the volume of the sphere(s) is not buried in the docked ligand.
- Click on **Descriptors** in the top-level menu, then **Define** in the next menu, followed by **Add Sphere Volume Descriptor**.
- Set the descriptor name to *W84 accessibility* (the descriptor will measure how much sphere volume is *unoccupied* by atoms).
- Hit the **Add New Sphere** button. This opens an *Edit Sphere* dialogue box and immediately creates (and displays in the visualiser) a sphere of 1.4Å radius at the origin (this may not be visible in the visualiser as the protein is centred on a point a long way from the origin). We need to move this sphere to a position close to Trp84, but, ideally, not overlapping too much with that residue.
- Switch on the *Centroid of atoms selected in 3D viewer* option and then pick atoms in and near to Trp84 (you can pick ligand as well as protein atoms) until the sphere is in a reasonable position.



- Hit **Done** to complete the sphere definition.
- Set the *Save Options* so that we get the percentage of sphere volume, not the absolute amount; then hit **Define**.

24.2.7 Saving the Descriptor Configuration

- Having set up several descriptors for a particular protein, you may want to calculate them repeatedly for successive docking runs. A descriptor configuration can be saved via the menu command **Descriptors** followed by **Export Descriptor Configuration File**. For tutorial purposes, a file containing the descriptors we have just defined, together with several others, is provided, so you do not have to save anything now.

You can safely quit Hermes at this point, or proceed to the final part of the tutorial which describes how to calculate the descriptors and gives some pointers on their use.

24.3Part 3: Calculating and Using Descriptors

If you closed Hermes after completing the previous part of this tutorial, re-open it and load in the tutorial GOLD .conf file (*gold_gs_top1.conf*).

- (See “Loading a GOLD Run” on page 122. for information on where to find this and all other files referred to below).

24.3.1 Loading in a Previously Defined Set of Descriptors

- Load in the descriptor configuration file provided with this tutorial by hitting the top-level **Descriptor** button followed by **Import Descriptor Configuration File** in the next menu. The name of the file to load is *hermes_tutorial_config_2.xml*.
- To view the list of defined descriptors click on the **Descriptors** tab in the *Protein Explorer*. The new set of descriptors contains all those you created earlier in the tutorial plus several others.
- In the *Descriptors* tab you can view details of any descriptor by double-clicking on it. This will display a dialogue box which shows the descriptor settings (and also allows you to edit them). Hit **Cancel** after you have inspected the settings, since we do not want to do any descriptor editing at this point.

24.3.2 The Descriptors

- This table explains the descriptors you have just read in. If you are not interested and just want to know how to calculate descriptors, skip to the next section.

Descriptor Name(s)	Meaning
<i>molec wt</i>	Ligand molecular weight.
<i>rot bond</i>	Number of rotatable bonds in ligand.
<i>bump/H-bond count</i>	Number of ligand atoms forming (a) H-bonds and (b) short contacts (sum of vdw radii - 0.5Å or shorter) to protein.
<i>very bad bumps</i>	Number of very bad protein-ligand clashes (sum of vdw radii - 1.0Å or shorter).
<i>polar-H count;</i> <i>acceptor count;</i> <i>occluded lig polar-H count;</i> <i>occluded lig acceptor count</i>	Number of polar H atoms in ligand; number of H-bond acceptor atoms in ligand; number of ligand polar H atoms that are not H-bonded to the protein and are also not solvent accessible; similar quantity for ligand acceptors.
<i>occluded prot polar-H count;</i> <i>occluded prot acceptor count</i>	Number of protein polar H atoms that are not H-bonded to the ligand and are also not solvent accessible; similar quantity for protein acceptors.

Descriptor Name(s)	Meaning
<i>buried total;</i> <i>buried hydrophobic</i>	Percentage of ligand surface area that remains solvent accessible after docking; percentage of hydrophobic ligand surface area accessible after docking.
<i>W84 accessibility</i>	Percentage volume of a sphere positioned near Trp84 that is unoccupied after ligand docks.
<i>W279 accessibility</i>	Percentage volume of a pair of spheres positioned near Trp279 that is unoccupied after ligand docks.
<i>S200 H-bond count;</i> <i>oxyanion H-bond count;</i> <i>oxyanion accessibility</i>	Number of H-bonds donated to ligand by hydroxyl hydrogen of Ser200; number of H-bonds donated to ligand by the three N-H groups that constitute the oxyanion hole; percentage volume of a pair of spheres positioned near Ser200 and the oxyanion hole that are unoccupied by ligand hydrophobic atoms.

24.3.3 Calculating Descriptors

- Click on **Descriptors** followed by **Calculate**.
- Select the option *Current GOLD docking solutions*.
- You now have two options regarding the descriptor calculation output. You can:
 - Output to a *Molecule file* (activate the adjacent tick box, browse to the directory where the file will be saved and specify the filename).
 - Save a *comma-separated values* file (activate the adjacent tick box, browse to the directory where the file will be saved and specify the filename).
- For GOLD Suite users: if a GoldMine DB had been created a further two options would be available:
 - In the *Input ligands* part of the dialogue we could select ligands from the GoldMine DB to calculate the descriptors for by selecting **GoldMine selection**.
 - In the *Output options* part of the dialogue we could choose to save the descriptors to **This GoldMine**. The calculated descriptors could then be inspected and/or combined with other descriptors or docking parameters using Boolean logic. See the GoldMine documentation for further details.
- Select the output option *Save comma-separated value file* and specify a suitable file name (the .csv extension will be added automatically).
- Hit **Run**. The calculation starts immediately and the output is displayed in a pop-up window.
- The calculation will take many minutes so, now that you have established the calculation will run properly, hit **Interrupt**.

- The results of the complete calculation are provided in the file *hermes_tutorial_descriptors.csv*. If you want to repeat the complete calculation yourself, click again on **Descriptors ... Calculate descriptors**. The dialogue will remember the settings from before. This time, hit **Submit & Exit**; this will run the calculation as a background process and close down Hermes, so that you can pick up your results later. You will not be able to open the output .csv file until the calculation is complete.

24.3.4 Descriptor Output

- The remainder of this tutorial works on the assumption that a licence is held only for Hermes.
- If you are a GOLD Suite licence holder and have opted to save the descriptor calculation results to a GoldMine, subsequent analysis can be carried out using GoldMine.
- The .csv file we have output can be analysed in a program of your choice. In the remainder of this tutorial, we use Microsoft Excel.
- Read the output file, *hermes_tutorial_descriptors.csv*, into Excel or some other program in which you can view the file easily.
- Most of the output is self-explanatory but here are some points to note:
 - The descriptor names are augmented by additional text (sometimes useful, sometimes not, often depending on how informative the descriptor names were in the first place).
 - The index column gives the position of each ligand in the input file; this is also displayed alongside each ligand name in the *Ligands* list box at the bottom-right of the Hermes graphical interface.
 - The *very_bad_bumps* descriptor results in several columns of output because the contacts are separated according to the protein atom that they involve. A similar comment applies to the *S200 H-bond count* and *oxyanion H-bonds* descriptors, and the columns are not necessarily in the order you might expect.
 - Zero values (e.g. for the descriptor *S200_H-bond_count* when the serine hydroxyl hydrogen forms no H-bonds) are sometimes reported as blanks rather than the number 0.
 - Two values are output for each of the sphere-volume descriptors, being the free volume of the sphere(s) after and before docking. If the sphere positioning is such that there is no overlap with the protein, the free volume before docking should always be 100%. If there is some overlap, the number will be < 100% and may vary from ligand to ligand (because GOLD may move some protein H atoms to optimise H-bonding).
 - Buried ligand surface area is given both after and before docking. The value before docking must always be 0 if (as in this case) the descriptor has been set up to report percentages rather than absolute surface areas.
 - The *bump/H-bond_count* descriptor results in two values, one for bumps, one for H-bonds.

Experienced drug designers will need no advice on how to use descriptors but, for completion, the

remainder of this tutorial describes a simple analysis that we performed with the descriptor data. (If you want to skip this, the tutorial is finished.)

24.3.5 Tidying up the Descriptor Output

- Open the file *hermes_tutorial_descriptors.xls* in Excel.
- To create this file from the original output, we have:
 - Summed the number of H-bonds donated by the three oxyanion N-H and the Ser200 O-H hydrogens. The result is in the column entitled *Number of H-bonds donated by Ser200 OH and oxyanion hole NHs*.
 - Summed the *very_bad_bumps* columns to give a total bad-bump count (column headed *Number of very bad protein-ligand clashes*).
 - Created two new columns derived from the Gold scoring-function terms. The first, *GOLD fitness score without internal terms*, is the total GOLD score minus the three terms (H-bonding, torsional, vdw) that model the internal energy of the ligand. Exclusion of the internal terms is known to improve correlation with binding affinity on occasions. The other column is the same quantity normalised by molecular weight.
 - Hidden or removed unwanted columns (e.g. the sphere-volume and surface-area values before docking).
 - Reordered several of the columns.
 - Done some tidying up, assigned more meaningful column names, etc.
- Visual scanning of the spreadsheet quickly highlights a number of ligands that can probably be eliminated, e.g.
 - those that form very bad clashes;
 - those that do not form H-bonds to the oxyanion-hole protein donor groups but which have hydrophobic atoms in this region of the active site (indicated by low values in the column headed *Percentage oxyanion hole sphere volume unoccupied by hydrophobic atoms*).
- Conversely, there are a several ligands that occupy space near to Trp84 or (much less commonly) Trp279, both of which are known to be important binding regions.

24.3.6 Filtering out Unwanted Ligands

- Next, load *hermes_tutorial_descriptors_filtered.xls* into Excel. This is the same spreadsheet as the one discussed above but now the ligands have been sorted on the column headed *External (protein-ligand) Goldscore term divided by molecular weight* and filters have been used to hide ligands meeting any of the following criteria, all of which are likely to be associated with poor binding:
 - There is at least 1 very bad clash (shorter than sum of vdw radii - 1.0Å).
 - There are >10 ligand atoms forming fairly bad clashes (shorter than sum of vdw radii - 0.5Å).
 - There are 2 or more occluded polar hydrogens on the ligand, i.e. polar hydrogens that are not

H-bonded to the protein and are not solvent accessible. Such atoms are probably in contact with hydrophobic protein atoms, which is clearly unfavourable. However, they could be forming intramolecular H-bonds to a ligand acceptor atom, which is why we allow up to one occluded polar H to occur.

- There are 2 or more occluded polar hydrogens on the protein.
- The Trp84 sphere has at least 20% accessibility (this eliminates any ligand that does not interact with this key binding residue).
- The buried ligand surface area is <95%.
- The oxyanion-hole sphere has <25% accessibility and the Ser200 hydroxyl group and oxyanion-hole NH groups form no hydrogen bonds (implying that ligand hydrophobic atoms are at preventing these donor groups from forming H-bonds).

24.3.7 Results

- Of the ligands that were not filtered out, the top-ranked ligand, number 145, is a very small molecule which is clearly not going to be of high affinity on its own. Nevertheless, it forms some nice-looking interactions and could maybe generate some good ideas for ligand design.
- The second in the ranking (index number 1) is a known AChE ligand.
- The other known ligand (index number 11) was filtered out by the oxyanion-hole sphere criterion. The reason is that GOLD docked this ligand incorrectly (the aromatic ring is the wrong way round; the phenol OH should be near the oxyanion hole). This emphasises that, if the docking pose is incorrect, there is little chance of predicting affinity.
- Load the GOLD run *gold_cs_top1.conf* into Hermes. This is the same set of ligands, but docked using the Chemscore function. (For some reason unknown to this tutorial writer, the entry identifiers of the ligands are different from the Goldscore dockings, but they are in the same order so the index numbers are still the same.)
- When the Chemscore docking function is used for docking, ligand 11 is docked the right way round.
- The files *hermes_tutorial_chemscore_descriptors.xls* and *hermes_tutorial_chemscore_descriptors_filtered.xls* contain, respectively, the tidied Hermes descriptor output for this set of dockings and the file produced when filters are applied as described above. The ligands that survived the filters are now ordered on Chemscore/mw. Both of the known inhibitors (1, 11) are now high in the rankings.

This ends the tutorial.

